REPORT RESUMES

ED 002 913 24

AN EMPIRICAL INVESTIGATION OF THE EFFECTS OF NONNORMALITY UPON THE SAMPLING DISTRIBUTION OF THE PROJECT MOMENT CORRELATION COEFFICIENT.

BY- NORRIS, RAYMOND C. HJELM, HOWARD
GEORGE PEABODY COLL. FOR TEACHERS, NASHVILLE, TENN.

REPORT NUMBER CRP-637 PUB DATE 15 MAR 60
EDRS PRICE MF-$0.27 HC-$7.20 180P.

DESCRIPTORS- POPULATION DISTRIBUTION, COMPARATIVE TESTING, *GROUP INTELLIGENCE TESTING, PSYCHOLOGICAL TESTING, *FACTOR ANALYSIS, *MEASUREMENT TECHNIQUES, EVALUATION, *POPULATION TRENDS, NASHVILLE, TENNESSEE

THE STUDY EMPIRICALLY DETERMINED THE EFFECTS OF NONNORMALITY UPON SOME SAMPLING DISTRIBUTIONS OF THE PRODUCT MOMENT CORRELATION COEFFICIENT (PMCC). SAMPLING DISTRIBUTIONS OF THE PMCC WERE OBTAINED BY DRAWING NUMEROUS SAMPLES FROM CONTROL AND EXPERIMENTAL POPULATIONS HAVING VARIOUS DEGREES OF NONNORMALITY AND BY CALCULATING CORRELATION COEFFICIENTS FOR THESE SAMPLES. THE SAMPLING DISTRIBUTIONS OBTAINED FROM THE EXPERIMENTAL POPULATIONS WERE COMPARED WITH THE THEORETICAL SAMPLING DISTRIBUTIONS OF PMCC OF RANDOM SAMPLES FROM BIVARIATE NORMAL POPULATIONS AND WITH THE SAMPLING DISTRIBUTIONS OBTAINED FROM THE CONTROL POPULATIONS WHICH WERE APPROXIMATELY BIVARIATE NORMAL. THE RESULTS OF THE INVESTIGATION OF THE EFFECTS OF NONNORMALITY UPON THE SAMPLING DISTRIBUTIONS OF THE PMCC INDICATED THAT NONNORMALITY HAD AN AFFECT AND THAT IT WAS OF SUCH A MAGNITUDE THAT IT SHOULD BE TAKEN INTO ACCOUNT IN DEALING WITH STATISTICAL TESTS OF INFERENCE INVOLVING PRODUCT MOMENT CORRELATION. MORE RESEARCH IS NEEDED, ESPECIALLY IN REGARD TO OTHER TYPES OF NONNORMALITY AND POPULATIONS NOT HAVING IDENTICAL MARGINAL FREQUENCIES. (GD)

ED002913

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

AN EMPIRICAL INVESTIGATION OF THE EFFECTS OF NONNORMALITY UPON THE SAMPLING DISTRIBUTIONS OF THE PRODUCT MOMENT CORRELATION COEFFICIENT

Cooperative Research Project No. 637

Raymond C. Norris

and

Howard F. Hjelm

George Peabody College for Teachers

March 15, 1960

The research reported herein was performed pursuant to a contract with the Office of Education, United States Department of Health, Education, and Welfare.

3414A
227

ACKNOWLEDGMENTS

The authors are indebted to Mr. Ira Grissom, applied scientist, and Miss Joanne Albright, machine operator, of The International Business Machine Corporation for their generous cooperation in preparing card populations and in checking out the program for the IBM 650 electronic calculator. Appreciation is also due Mr. Edward J. Boling, Director of the Budget for the State of Tennessee, Mr. Harlan Mathews, Assistant Director of the Budget, and Mr. Edward Criddle, machine operator, for making available the IBM 650 Data Processing System of the State of Tennessee. Dr. William F. Atchison, Director of the Rich Computer Center, Georgia Institute of Technology, also generously made facilities available for partial processing of data for this study.

The research reported herein was performed under Cooperative Research Project No. 637, Office of Education, United States Department of Health, Education, and Welfare.

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS ii

LIST OF TABLES v

LIST OF GRAPHS xii

CHAPTER

I INTRODUCTION. 1

II RELATED RESEARCH. 15

Egon S. Pearson's Study: 1929. 15
Hilda Frost Dunlap's Study. 18
George A. Baker's Study 20
Egon S. Pearson's Study: 1931. 23
Egon S. Pearson's Study: 1932. 28
Leone Chesire, Elena Oldis, and Egon S. Pearson's Study 32
Paul R. Rider's Study 39
Conclusions 43

III PROCEDURE 44

IV RESULTS AND CONCLUSIONS 55

Populations Having Approximately No Correlation . . . 56
Normal (Control) Population. 56
Rectangular Population. 60
Leptokurtic Population. 61
Skewed Populations. 62
Populations Having Substantial Correlations 63
Normal (Control) Population 64
Rectangular Population. 67
Leptokurtic Population. 68
Skewed Populations. 69
General Conclusions 70
Relationship to Findings in Previous Research 72

V SUMMARY 74

Procedure 74
Results and Conclusions 76

BIBLIOGRAPHY 80

Page

APPENDIX. 83

A TABLES OF BIVARIATE POPULATIONS. 84

B TABLES OF SAMPLING DISTRIBUTIONS OF THE PRODUCT MOMENT CORRELATION COEFFICIENT 95

C GRAPHS OF SAMPLING DISTRIBUTIONS OF THE PRODUCT MOMENT CORRELATION COEFFICIENT 133

LIST OF TABLES

Table Page

1 An Example of a Nonnormal Bivariate Distribution Having Normal Univariate Distributions 7

2 Analysis of Sums of Squares 10

3 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient (ρ = +.5346, N = 20) from First Experiment of Egon S. Pearson's Study: 1929. 16

4 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient (ρ = +.4626, N = 20) from the Second Experiment of Egon S. Pearson's Study: 1929. 18

5 Proportional Frequencies for the Marginal Distributions for the Skewed Population Used in George A. Baker's Study 21

6 Obtained Correlation Coefficients from George A. Baker's Study. 22

7 Grouped Sampling Distributions of Product Moment Correlation Coefficients from George A. Baker's Study 23

8 Obtained and Theoretical Sampling Distributions of Product Moment Correlation Coefficients Together with Tests of Goodness of Fit (ρ = 0, N = 10) from Egon S. Pearson's Study: 1931. 25

9 Obtained and Theoretical Sampling Distributions of Product Moment Correlation Coefficients Together with Tests of Goodness of Fit (ρ = 0, N = 20) from Egon S. Pearson's Study: 1931. 26

10 Obtained and Theoretical Standard Errors of the Product Moment Correlation Coefficient (ρ = 0, N = 10, N = 20) from Egon S. Pearson's Study: 1931. 27

Table Page

11 Obtained and Theoretical Sampling Distributions of Product Moment Correlation Coefficients Together with Tests of Goodness of Fit ($\rho = 0$, $N = 10$) from Egon S. Pearson's Study: 1932. 29

12 Obtained and Theoretical Sampling Distributions of Product Moment Correlation Coefficients Together with Tests of Goodness of Fit ($\rho = 0$, $N = 20$) from Egon S. Pearson's Study: 1932. 30

13 Obtained and Theoretical Standard Errors of the Product Moment Correlation Coefficient ($\rho = 0$, $N = 10$, $N = 20$) from Egon S. Pearson's Study: 1932. 31

14 First Nonnormal Population Having a Correlation of .50 from Leone Chesire, Elena Oldis, and Egon S. Pearson's Study 33

15 Second Nonnormal Population Having a Correlation of .5462 from Leone Chesire, Elena Oldis, and Egon S. Pearson's Study 34

16 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient for Sample Size Five from Leone Chesire, Elena Oldis, and Egon S. Pearson's Study. 36

17 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient for Sample Size Ten from Leone Chesire, Elena Oldis, and Egon S. Pearson's Study. 37

18 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient for Sample Size Twenty from Leone Chesire, Elena Oldis, and Egon S. Pearson's Study. 38

19 Obtained and Theoretical Mean Product Moment Correlation Coefficients, Standard Errors of the Product Moment Correlation Coefficient, and Results of Tests of Goodness of Fit of Obtained and Theoretical Sampling Distributions of Product Moment Correlation Coefficients from Leone Chesire, Elena Oldis, and Egon S. Pearson's Study 39

Table | Page

20 Obtained and Theoretical Sampling Distributions from Paul R. Rider's Study. 41

21 Obtained and Theoretical Mean Correlation Coefficients and Standard Errors of the Correlation Coefficient from Paul R. Rider's Study 42

22 Characteristics of Univariate Distributions Used. 46

23 Correlation Coefficients of the Established Populations 47

24 Fields on Master Cards and Detail Cards 50

25 Correlations of Established Populations and of Card Populations. 54

26 Results of Kolmogorov-Smirnov Test Comparing Obtained Sampling Distributions from Populations Having Approximately No Correlation with Their Theoretical Sampling Distributions 57

27 Proportions of Significant Coefficients in Sampling Distributions from Populations Having Approximately No Correlation ($\alpha = .05$) 58

28 Proportions of Significant Coefficients in Sampling Distributions from Populations Having Approximately No Correlation ($\alpha = .01$) 59

29 Results of Kolmogorov-Smirnov Tests Comparing Obtained Sampling Distributions from Populations Having Substantial Correlation with Their Theoretical Sampling Distributions 64

30 Proportions of Significant Coefficients in Sampling Distributions from Populations Having Substantial Correlation ($\alpha = .05$). 65

31 Proportions of Significant Coefficients in Sampling Distributions from Populations Having Substantial Correlation ($\alpha = .01$). 66

Table Page

32 Bivariate Normal Population Having a Correlation of .000 85

33 Bivariate Rectangular Population Having a Correlation of .000 86

34 Bivariate Leptokurtic Population Having a Correlation of .000 87

35 Bivariate Slightly Skewed Population Having a Correlation of -.005. 88

36 Bivariate Markedly Skewed Population Having a Correlation of -.004. 89

37 Bivariate Normal Population Having a Correlation of +.832. 90

38 Bivariate Rectangular Population Having a Correlation of +.834. 91

39 Bivariate Leptokurtic Population Having a Correlation of +.813. 92

40 Bivariate Slightly Skewed Population Having a Correlation of +.828. 93

41 Bivariate Markedly Skewed Population Having a Correlation of +.832. 94

42 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Normal Population (ρ = +.016, N = 15). 96

43 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Rectangular Population (ρ = -.019, N = 15) . . 98

44 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Leptokurtic Population (ρ = -.004, N = 15) . . 100

45 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Slightly Skewed Population (ρ = -.010, N = 15) 102

Table Page

46 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Markedly Skewed Population (ρ = -.012, N = 15) 104

47 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Normal Population (ρ = +.016, N = 30). 106

48 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Rectangular Population (ρ = -.019, N = 30) . . 107

49 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Leptokurtic Population (ρ = -.004, N = 30) 108

50 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Slightly Skewed Population (ρ = -.010, N = 30) 109

51 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Markedly Skewed Population (ρ = -.012, N = 30) 110

52 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Normal Population (ρ = +.016, N = 90). 111

53 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Rectangular Population (ρ = -.019, N = 90) . . 112

54 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Leptokurtic Population (ρ = -.004, N = 90) . . 113

55 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Slightly Skewed Population (ρ = -.010, N = 90) 114

56 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Markedly Skewed Population (ρ = -.012, N = 90) 115

Table Page

57 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Normal Population (ρ = +.835, N = 15). 116

58 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Rectangular Population (ρ = +.831 N = 15). . . 117

59 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Leptokurtic Population (ρ = +.809, N = 15) . . 118

60 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Slightly Skewed Population (ρ = +.826, N = 15) 120

61 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Markedly Skewed Population (ρ = +.826, N = 15) 121

62 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Normal Population (ρ = +.835, N = 30). 122

63 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Rectangular Population (ρ = +.831, N = 30) . . 123

64 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Leptokurtic Population (ρ = +.809, N = 30) . . 124

65 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Slightly Skewed Population (ρ = +.826, N = 30) 125

66 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Markedly Skewed Population (ρ = +.826, N = 30) 126

67 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Normal Population (ρ = +.835, N = 90). 127

Table Page

68 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Rectangular Population (ρ = +.831, N = 90) . . 128

69 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Leptokurtic Population (ρ = +.809, N = 90) . . 129

70 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Slightly Skewed Population (ρ = +.826, N = 90) 131

71 Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Markedly Skewed Population (ρ = +.826, N = 90) 132

LIST OF GRAPHS

Graph		Page
1	Distribution of 15,984 Correlation Coefficients Based on Normal Population (ρ = +.016, N = 15)	134
2	Distribution of 15,984 Correlation Coefficients Based on Rectangular Population (ρ = -.019, N = 15). . .	135
3	Distribution of 15,984 Correlation Coefficients Based on Leptokurtic Population (ρ = -.004, N = 15). . .	136
4	Distribution of 15,984 Correlation Coefficients Based on Slightly Skewed Population (ρ = -.010, N = 15).	137
5	Distribution of 15,984 Correlation Coefficients Based on Markedly Skewed Population (ρ = -.012, N = 15).	138
6	Distribution of 7,992 Correlation Coefficients Based on Normal Population (ρ = +.016, N = 30)	139
7	Distribution of 7,992 Correlation Coefficients Based on Rectangular Population (ρ = -.019, N = 30).	140
8	Distribution of 7,992 Correlation Coefficients Based on Leptokurtic Population (ρ = -.004, N = 30).	141
9	Distribution of 7,992 Correlation Coefficients Based on Slightly Skewed Population (ρ = -.010, N = 30). . . .	142
10	Distribution of 7,992 Correlation Coefficients Based on Markedly Skewed Population (ρ = -.012, N = 30). . . .	143
11	Distribution of 2,664 Correlation Coefficients Based on Normal Population (ρ = +.016, N = 90)	144
12	Distribution of 2,664 Correlation Coefficients Based on Rectangular Population (ρ = -.019, N = 90).	145
13	Distribution of 2,664 Correlation Coefficients Based on Leptokurtic Population (ρ = -.004, N = 90).	146
14	Distribution of 2,664 Correlation Coefficients Based on Slightly Skewed Population (ρ = -.010, N = 90). . . .	147
15	Distribution of 2,664 Correlation Coefficients Based on Markedly Skewed Population (ρ = -.012, N = 90). . . .	148

Graph Page

16 Distribution of 15,984 Correlation Coefficients Based on Normal Population (ρ = +.835, N = 15). 149

17 Distribution of 15,984 Correlation Coefficients Based on Rectangular Population (ρ = +.831, N = 15) 150

18 Distribution of 15,984 Correlation Coefficients Based on Leptokurtic Population (ρ = +.809, N = 15) 151

19 Distribution of 15,984 Correlation Coefficients Based on Slightly Skewed Population (ρ = +.826, N = 15) . . . 152

20 Distribution of 15,984 Correlation Coefficients Based on Markedly Skewed Population (ρ = +.826, N = 15) . . . 153

21 Distribution of 7,992 Correlation Coefficients Based on Normal Population (ρ = +.835, N = 30). 154

22 Distribution of 7,992 Correlation Coefficients Based on Rectangular Population (ρ = +.831, N = 30) 155

23 Distribution of 7,992 Correlation Coefficients Based on Leptokurtic Population (ρ = +.809, N = 30) 156

24 Distribution of 7,992 Correlation Coefficients Based on Slightly Skewed Population (ρ = +.826, N = 30) . . . 157

25 Distribution of 7,992 Correlation Coefficients Based on Markedly Skewed Population (ρ = +.826, N = 30) . . . 158

26 Distribution of 2,664 Correlation Coefficients Based on Normal Population (ρ = +.835, N = 90). 159

27 Distribution of 2,664 Correlation Coefficients Based on Rectangular Population (ρ = +.831, N = 90) 160

28 Distribution of 2,664 Correlation Coefficients Based on Leptokurtic Population (ρ = +.809, N = 90) 161

29 Distribution of 2,664 Correlation Coefficients Based on Slightly Skewed Population (ρ = +.826, N = 90) . . . 162

30 Distribution of 2,664 Correlation Coefficients Based on Markedly Skewed Population (ρ = +.826, N = 90) . . . 163

CHAPTER I

INTRODUCTION

The product moment correlation coefficient is the familiar index of linear relationship commonly known as Pearson's r. Its origin is generally traced back to Galton and Karl Pearson.[1]

Confusion exists in the literature in regard to the assumptions necessary in dealing with product moment correlation. The fact that the significance of the underlying assumption of normality is in dispute is pointed out by three recent articles in The American Psychologist.[2, 3, 4] Much of the confusion seems to be due to not explicitly stating which concept (s) of product moment correlation is/are being considered and also whether or not both variables are considered independent or only one variable is considered independent and the other variable is considered dependent in tests of significance.

The formula for the product moment correlation coefficient may be derived mathematically using the method of least squares and the assumption that the bivariates are linearly related.[5, 6]

---

[1] Karl Pearson, "Notes on the History of Correlation," Biometrika, XIII, No. 1 (1920), 25-45.

[2] M. D. Nefsger and James Drasgow, "The Needless Assumption of Normality in Pearson's r," The American Psychologist, XII, No. 10 (1957), 623-625.

[3] "Comment on 'The Needless Assumption of Normality in Pearson's r,'" The American Psychologist, XIII, No. 9 (1958), 544-546.

[4] Arnold Binder, "Considerations of the Place of Assumptions in Correlational Analysis," The American Psychologist, XIV, No. 8 (1959), 504-510.

[5] Maurice G. Kendall, The Advanced Theory of Statistics, I (3rd ed.; London: Charles Griffin & Company Limited, 1947), pp. 327-330.

[6] Charles C. Peters and Walter R. Van Voorhis, Statistical

x and y are deviation values of X and Y from their respective means.

Assume that y is linearly related to x.

$\tilde{y}$ is a regression estimate of y.

$\tilde{y} = b_{yx}x + a_{yx}$ because y is linearly related to x. (1)

e is deviation of y from its regression estimate $\tilde{y}$.

$\Sigma e^2 = \Sigma(y - \tilde{y})^2$ (2)

Substituting equation (1) into equation (2),

$\Sigma e^2 = \Sigma(y - b_{yx}x - a_{yx})^2$ (3)

Applying the method of least squares,

$\partial \Sigma e^2 / \partial b_{yx} = \Sigma 2(y - b_{yx}x - a_{yx})(-x) \longrightarrow$ a min.

Setting the partial equal to zero and solving for b_{yx},

$b_{yx} = (\Sigma xy - a_{yx}\Sigma x)/\Sigma x^2$

$\Sigma x = 0$ because x's are deviation values.

$\therefore b_{yx} = \Sigma xy/\Sigma x^2$

Similarly,

$b_{xy} = \Sigma xy/\Sigma y^2$

By definition the product moment correlation coefficient is equal to the geometric mean of the regression coefficients b_{yx} and b_{xy}.

$$r = (b_{yx}b_{xy})^{1/2}$$
$$= ((\Sigma xy/\Sigma x^2)(\Sigma xy/\Sigma y^2))^{1/2}$$
$$= \Sigma xy/(\Sigma x^2 \Sigma y^2)^{1/2}$$

Procedures and Their Mathematical Bases (1st ed.; New York: McGraw-Hill Book Company, Inc., 1940), pp. 91-101.

The regression coefficients b_{yx} and b_{xy} were derived in the derivation of the product moment correlation coefficient. The only assumption made in the derivation of b_{yx} was that Y was linearly related to X. Similarly, the only assumption made in the derivation of b_{xy} was that X was linearly related to Y. The assumption of linearity should be met if a regression coefficient is to be used to predict one variable from a knowledge of another variable. No other assumption was made in the derivation of the formula for the product moment correlation coefficient. While the assumption of linearity need not be met in order to calculate the product moment correlation coefficient of a distribution of bivariates, it seems reasonable that the assumption should be met if the obtained coefficient is to be interpreted as an index of relationship between X and Y.

The standard error of estimate is a measure of the variability of the variates about the regression line. The measure of variability of the Y values about the regression line $b_{yx} + a_{yx}$ is $s_{y.x}$ and the measure of the variability of the X values about the regression line $b_{xy} + a_{xy}$ is $s_{x.y}$. The assumptions entering into the derivations of the standard errors of estimate are that Y is linearly related to X for $s_{y.x}$ and that X is linearly related to Y for $s_{x.y}$. The assumption of homoscedasticity, that the Y arrays are homoscedastic for $s_{y.x}$ and that the X arrays are homoscedastic for $s_{x.y}$, is logical if a single measure of variability is to be obtained for expressing the variability of a variable about a regression line. If the standard errors of estimate are to be used in establishing confidence statements, it is implied that the Y arrays are normally distributed in the interpretation of $s_{y.x}$ and that the X arrays are normally distributed in the interpretation of $s_{x.y}$.

Assuming the assumptions which entered into the derivations of b_{yx}, b_{xy}, and r have been met, the formulae for the standard errors of estimate may be derived algebraically.[7]

x and y are deviation values of X and Y from their respective means.

Assume that y is linearly related to X.

$\tilde{y}$ is a regression estimate of y.

e is deviation of y from its regression estimate $\tilde{y}$.

N is the bivariate frequency.

$$s_{y.x}^2 = \Sigma e^2/N$$

$$= \Sigma (y - \tilde{y})^2/N$$

$$= \Sigma y^2/N - 2\Sigma y\tilde{y}/N + \Sigma \tilde{y}^2/N \qquad (4)$$

Applying the method of least squares to equation (3),

$$\partial \Sigma e^2/\partial a_{yx} = \Sigma 2(y - b_{yx} - a_{yx})(-1) \longrightarrow \text{a min.}$$

Setting the partial equal to zero and solving for a_{yx},

$$a_{yx} = (b_{yx}\Sigma x - \Sigma y)/N$$

$\Sigma y = 0$ because y's are deviation values.

$\Sigma x = 0$ because x's are deviation values.

$$\therefore a_{yx} = 0 \qquad (5)$$

Substituting equation (5) into equation (1),

---

[7] Quinn McNemar, Psychological Statistics (2nd ed.; New York: John Wiley & Sons, Inc., 1955), pp. 131-135.

$$\begin{aligned}\tilde{y} &= b_{yx}x \\ &= (\Sigma\, xy/\Sigma\, x^2)x \\ &= (\Sigma\, xy/(\Sigma\, x^2 \Sigma\, y^2)^{1/2})((\Sigma\, y^2/N)^{1/2}/(\Sigma\, x^2/N)^{1/2})x \\ &= r(s_y/s_x)x \qquad (6)\end{aligned}$$

Substituting equation (6) into equation (4) and simplifying,

$$\begin{aligned}s_{y.x}^{\ 2} &= s_y^{\ 2} - 2r(s_y/s_x)\,\Sigma\, xy/N + r^2(s_y^{\ 2}/s_x^{\ 2})\,\Sigma\, x^2/N \\ &= s_y^{\ 2} - r^2 s_y^{\ 2} \\ &= s_y^{\ 2}(1 - r^2) \\ s_{y.x} &= s_y(1 - r^2)^{1/2}\end{aligned}$$

Similarly,

$$s_{x.y} = s_x(1 - r^2)^{1/2}$$

The coefficient of determination is defined as the proportion of the variance of Y which can be attributed to variation in X. This is equivalent to the variance of the weighted predicted values for Y divided by the total variance of Y.

x and y are deviation values of X and Y from their respective means.

$\tilde{y}$ is a regression estimate of y.

N is the bivariate frequency.

$$\begin{aligned}\text{Coefficient of determination} &= (\Sigma\,\tilde{y}^2/N)/(\Sigma\, y^2/N) \\ &= \Sigma\,\tilde{y}^2/\Sigma\, y^2 \\ &= b_{yx}^{\ 2}\,\Sigma\, x^2/\Sigma\, y^2 \\ &= (\Sigma\, xy/\Sigma\, x^2)^2\,\Sigma\, x^2/\Sigma\, y^2 \\ &= (\Sigma\, xy)^2/\Sigma\, x^2\,\Sigma\, y^2 \\ &= r^2\end{aligned}$$

Similarly, the coefficient of determination for the proportion of the variance of X which can be attributed to variation in Y is equal to r^2.

The only assumption entering into the derivation of the coefficient of determination is that of linearity.

The sampling distributions of the product moment correlation coefficient have been investigated mainly for samples randomly drawn from bivariate normal populations. The historical papers of Karl Pearson and R. A. Fisher deal with the sampling distributions of the product moment correlation coefficient from bivariate normal populations.

The mathematical equation for the bivariate normal surface is

$$z = \frac{1}{2\sigma x\sigma y(1-\rho^2)^{1/2}} e^{-\frac{1}{2(1-\rho^2)}(z_x^2 - 2\rho z_x z_y + z_y^2)}$$

where z equals the height of the surface above the XY plane, σx, σy, and ρ are parameters and z_x and z_y are X and Y values in standard units.

A bivariate normal surface has the following characteristics:

1. The regression is rectilinear.
2. The arrays are normally distributed.
3. The arrays are homoscedastic.
4. The arrays are homokurtic.
5. The arrays are homoclitic.
6. The marginal univariate distributions are normally distributed.
7. The horizontal slices are ellipses.

A necessary condition for a bivariate distribution to be normal is that the two univariate distributions be normal. However, the two univariate distributions can be normal without the bivariate distribution being normal.

The univariate distributions in Table 1 approximate the normal while the bivariate distribution does not.

Table 1

An Example of a Nonnormal Bivariate Distribution
Having Normal Univariate Distributions

Y Value	X Value 1	2	3	4	5	Total
5					1	1
4		2		2		4
3		2	3	1		6
2			3	1		4
1	1					1
Total	1	4	6	4	1	

F. N. David[8] constructed tables giving the areas and ordinates under the curves of various sampling distributions of the product moment correlation coefficient. The values in the tables are for product moment correlation coefficients for random samples which have been drawn from bivariate normal populations. The tables of areas and ordinates are given for values of r from -1.00 by intervals of .05 to +1.00 with finer intervals in places for sampling distributions for ρ of .0 by tenths to .9 for sample sizes 3 through 25, 50, 100, 200, and 400.

---

[8] F. N. David, Tables of the Ordinates and Probability Integral of the Distribution of the Correlation Coefficient in Small Samples (Cambridge, England: The University Press, 1938).

The standard error of the product moment correlation coefficient is equal to $(1 - \rho^2)/(N - 1)^{1/2}$, where N is the sample size. This formula for the standard error is based on the assumption that the population is bivariate normal. The above formula for the standard error of the product moment correlation coefficient is satisfactory if the sample size is large and if the correlation of the population is zero.[9]

The significance of the product moment correlation coefficient when its parameter is equal to zero may be tested by means of the t-ratio,

$$t = r(N - 2)^{1/2}/(1 - r^2)^{1/2}$$

with degrees of freedom equal to N - 2; or the F-ratio,

$$F = \frac{r^2/1}{(1 - r^2)/(N - 2)}$$

with degrees of freedom equal to 1 for the numerator and N - 2 for the denominator. The above tests are based on the assumption that the population is bivariate normal.[10]

A bivariate distribution is analogous to a distribution in the simple randomized design in one-way analysis of variance if one of the variables is considered independent and the other variable is considered dependent.

---

[9] Helen M. Walker and Joseph Lev, Statistical Inference (New York: Henry Holt and Company, 1953), pp. 251-252.

[10] Maurice G. Kendall, The Advanced Theory of Statistics, I (3rd ed.; London: Charles Griffin & Company Limited, 1947), pp. 339-343.

The independent arrays can be thought of as a series of populations.[11, 12]

The product moment correlation coefficient is an index measuring the closeness of fit of variates about the regression line. The total sum of squares of the deviations of Y from the mean of Y can be divided into two parts, the sum of squares of the deviations of Y from the regression estimates of Y and the sum of squares of the weighted deviations of regression estimates of Y from the mean of Y. Below is a derivation of the mean sums of squares for use in an F-ratio for testing the significance of r when the correlation of the population is zero and when X is considered the independent variable and Y is considered the dependent variable.

$y = (y - \mathcal{Y}) + \mathcal{Y}$, where y is the deviation of Y from the mean of Y and $\mathcal{Y}$ is the deviation of the regression estimate of Y from the mean of Y.

$$\begin{aligned} \Sigma y^2 &= \Sigma\,((y - \mathcal{Y}) + \mathcal{Y})^2 \\ &= \Sigma\,(y - \mathcal{Y})^2 + 2\Sigma\mathcal{Y}(y - \mathcal{Y}) + \Sigma\mathcal{Y}^2 \end{aligned} \qquad (7)$$

$$\begin{aligned} \Sigma\mathcal{Y}(y - \mathcal{Y}) &= \Sigma\, b_{yx}x(y - b_{yx}x) \\ &= \Sigma\, r(s_y/s_x)x(y - r(s_y/s_x)x) \\ &= r(s_y/s_x)\,\Sigma xy - r^2({s_y}^2/{s_x}^2)\,\Sigma x^2 \\ &= Nr^2{s_y}^2 - Nr^2{s_y}^2 \\ &= 0 \end{aligned} \qquad (8)$$

Substituting equation (8) into equation (7),

$$\Sigma y^2 = \Sigma\,(y - \mathcal{Y})^2 + \Sigma\mathcal{Y}^2 \qquad (9)$$

$$\begin{aligned} \Sigma\mathcal{Y}^2 &= {b_{yx}}^2\Sigma\, x^2 \\ &= (\Sigma\, xy/\Sigma\, x^2)^2\Sigma\, x^2 \\ &= r^2\Sigma\, y^2 \end{aligned} \qquad (10)$$

---

[11] Jerome C. R. Li, Introduction to Statistical Inference (Ann Arbor, Michigan: Edwards Brothers, Inc., 1957), pp. 244-266, 290-295.

[12] Quinn McNemar, Psychological Statistics (2nd ed.; New York:

Substituting equation (10) into equation (9) and solving for $\Sigma\,(y - \hat{y})^2$,

$$\Sigma(y - \hat{y})^2 = (1 - r^2)\Sigma\, y^2 \qquad (11)$$

Substituting equations (10) and (11) into equation (9),

$$\Sigma\, y^2 = (1 - r^2)\Sigma\, y^2 + r^2\Sigma\, y^2$$

The degrees of freedom for $\Sigma\, y^2$ is the total number of variates minus one $(N - 1)$. The degrees of freedom for $\Sigma\hat{y}^2$ is equal to one. And the degrees of freedom for $\Sigma(y - \hat{y})^2$ is equal to the total number of variates minus two $(N - 2)$.

Table 2

Analysis of Sums and Squares

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Sum of Squares
Regression	1	$r^2\,\Sigma y^2$	$r^2\,\Sigma y^2$
About Regression	$N - 2$	$(1 - r^2)\,\Sigma y^2$	$(1 - r^2)\,\Sigma y^2/(N - 2)$
Total	$N - 1$	$\Sigma\, y^2$	

The ratio of the mean sum of squares of regression over the mean sum of squares about regression is distributed as F with degrees of freedom of 1 for the numerator and $N - 2$ for the denominator when the correlation of the population is zero.

John Wiley & Sons, Inc., 1955), pp. 264-268.

$$F = \frac{\text{mean sum of squares of regression}}{\text{mean sum of squares about regression}}$$

$$= \frac{r^2 \Sigma y^2/1}{(1 - r^2)\, \Sigma y^2 (N - 2)}$$

$$= \frac{r^2/1}{(1 - r^2)/(N - 2)}, \qquad n_1 = 1,\ n_2 = N - 2$$

When the degrees of freedom for the numerator of an F-ratio is equal to one, the square root of F is distributed as t with the degrees of freedom for t equal to the degrees of freedom for the denominator of the F-ratio.

$$\therefore t = r(N - 2)^{1/2}/(1 - r^2)^{1/2}, \qquad n = N - 2$$

The assumptions entering into the test of the F-ratio for the significance of the correlation coefficient are that the variates have been drawn randomly, that the regression of the dependent variable on the independent variable is linear, that the dependent variable is normally distributed from array to array, that the dependent variable is homoscedastic from array to array, and that the correlation of the population is zero. No assumptions need be made in regard to the distribution of the independent variable. The above assumptions are analogous to the assumptions in the simple randomized design in one-way analysis of variance.

The sampling distributions of the product moment correlation coefficient are not symmetrical when their parameters are not equal to zero. R. A. Fisher[13, 14] developed a technique which is applicable for testing

---

[13] R. A. Fisher, "On the 'Probable Error' of a Coefficient of Correlation Deduced from a Small Sample," *Metron*, I, No. 4 (1921), 3-32.

[14] F. N. David, *Tables of the Ordinates and Probability Integral of the Distribution of the Correlation Coefficient in Small Samples* (Cambridge, England: The University Press, 1938), pp. xxxi-xxxiv.

the significance of a product moment correlation coefficient when its parameter is not equal to zero. Fisher applied a transformation to r and to ρ.

Fisher's transformations of r and ρ are

$$z_r = (1/2)\log_e((1+r)/(1-r))$$

and

$$\zeta = (1/2)\log_e((1+\rho)/(1-\rho))$$

z_r is approximately normally distributed about its mean

$$\bar{z}_r = \zeta + \frac{\rho}{2(N-1)}\left(1 + \frac{1+\rho^2}{8(N-1)} + \ldots\right) \qquad (12)$$

with a variance of

$$\sigma z_r^2 = \frac{1}{N-1}\left(1 + \frac{4-\rho^2}{2(N-1)} + \frac{176 - 21\rho^2 - 21\rho^4}{48(N-1)^2} + \ldots\right) \qquad (13)$$

The sampling distribution of z_r becomes a closer approximation to the normal curve as the sample size increases and as the absolute value of the correlation coefficient of the population decreases. Fisher's transformations are based on the assumption that the sample has been randomly drawn from a bivariate normal population.

Approximations of equations (12) and (13) are

$$\bar{z}_r = \zeta + \frac{\rho}{2(N-1)} \qquad (14)$$

and

$$\sigma z_r^2 = 1/(N-3) \qquad (15)$$

A further approximation of equation (12) is

$$\bar{z}_r = \zeta \qquad (16)$$

Equations (12) and (13) are the most accurate approximations of the above to the mean and variance of the sampling distribution of z_r. Equation (14) is a more accurate estimate than equation (16) of the mean of the sampling distribution of z_r. Equations (15) and (16) are the commonly used approximations in testing the significance of z_r.

The necessary assumptions for product moment correlation in descriptive statistics are the assumptions necessary in the derivation and interpretation of the statistics. The only assumption necessary in the derivation of the product moment correlation coefficient as a descriptive statistic is that of linearity. The assumption of one variable being linearly related to the other variable is also needed in the derivation of a standard error of estimate and the assumptions of homoscedasticity from array to array and normality of arrays are needed in its interpretation. The assumptions of homoscedasticity and normality are necessary only for the Y arrays for $s_{y.x}$ and only for the X arrays for $s_{x.y}$. Only the assumption of linearity of one variable on the other is necessary in the derivation and interpretation of a coefficient of determination.

The necessary assumptions in product moment correlation in statistical inference are the assumptions necessary for the statistic plus the assumptions necessary for the derivation of the sampling distributions for the statistical tests of significance. The sampling distributions of the product moment correlation coefficient determined by the standard error of the product moment correlation coefficient and of Fisher's transformations of r and ρ are based on the assumption that the populations are bivariate normal.

The assumptions necessary in the tests of significance of product moment correlation by means of the t-ratio $r(N - 2)^{1/2}/(1 - r^2)^{1/2}$ or the F-ratio $r^2(N - 2)/(1 - r^2)$ depend upon whether both the variables are considered independent or one of the variables is considered independent and the other variable is considered dependent. If both variables are independent the population is assumed to be bivariate normal. If one of the variables is independent and the other variable is dependent, the assumptions are not as strict as in the case when both variables are independent. In the case of one independent variable and one dependent variable the assumptions are that the dependent variable is linearly related to the independent variable, the dependent arrays are homoscedastic, and the dependent arrays are normal.

The purpose of the study was to determine by empirical means the effects of nonnormality upon some sampling distributions of the product moment correlation coefficient. Sampling distributions of product moment correlation coefficients were obtained by drawing numerous samples from control and experimental populations having various degrees of nonnormality and calculating correlation coefficients for these samples. The obtained sampling distributions from the experimental populations were compared with the theoretical sampling distributions of product moment correlation coefficients of random samples from bivariate normal populations and with the obtained sampling distributions from the control populations which were approximately bivariate normal populations.

CHAPTER II

RELATED RESEARCH

The following abstracts are of investigations of the effects of nonnormality upon the sampling distributions of the product moment correlation coefficient.

Egon S. Pearson's Study: 1929

Egon S. Pearson[1] conducted an investigation on the sampling distributions of the product moment correlation coefficient obtained from samples from two nonnormal populations. The sampling was done by using a table of random numbers.

In the first experiment the population approximated normality with regards to one variable but not with the other. The regression of Y on X was linear. The Y arrays were normally distributed and homoscedastic. There were three X arrays with marginal proportional frequencies of .40, .35, and .25. The correlation of the population was +.5346.

Two hundred samples of size twenty were randomly drawn from the first population. The distribution of obtained coefficients together with their theoretical frequencies are listed in Table 3. The mean obtained correlation coefficient was .5160, whereas the expected mean was .5244. The obtained standard error was .1614, whereas the expected standard error was .1704. Both of these obtained values were not considered significantly different from the expected values. A goodness of fit test comparing the obtained and theoretical distributions of correlation coefficients using eleven groups yielded a probability of .223.

---

[1] Egon S. Pearson, "Some Notes on Sampling Tests with Two Variables," Biometrika, XXI (1929), 337-360.

Table 3

Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient ($\rho = +.5346$, N = 20) from First Experiment of Egon S. Pearson's Study: 1929

r	Obtained Frequency	Theoretical Frequency
+.85	2	2.1 (+.85 and above)
+.80	1	6.0
+.75	13	12.3
+.70	21	18.5
+.65	26	22.8
+.60	19	24.5
+.55	21	23.8
+.50	15	21.4
+.45	19	18.0
+.40	24	14.4
+.35	15	11.1
+.30	6	8.2
+.25	8	5.8
+.20	5	4.0
+.15	3	2.7
+.10	1	1.8
+.05	0	1.1
.00	0	.7
-.05	1	(-.05 and below) .8
Total	200	200.0

In the second experiment the population was more nonnormal than the one in the first experiment. The regression was linear. The Y arrays were skewed and heteroscedastic. There were five X arrays with marginal proportional frequencies of .1667, .2666, .2167, .1833, and .1667 with standard deviations of 5.7471, 5.3191, 5.0000, 4.7619, and 4.5872, respectively. The correlation of the population was +.4626.

Three hundred samples of size thirty were randomly drawn from the second population. The distribution of obtained coefficients together with their theoretical frequencies are listed in Table 4. The mean obtained correlation coefficient was .4631, whereas the expected mean was .4563. The obtained standard error was .1475, whereas the expected standard error was .1488. Both of these obtained values were considered not significantly different from the expected values. A goodness of fit test comparing the obtained and theoretical distributions of correlation coefficients using fourteen groups yielded a probability of .916.

Pearson concluded that the results suggest that the normal bivariate surface can be distorted without seriously affecting the sampling distributions of the product moment correlation coefficient in samples of size as small as twenty. The X distributions in both populations were platykurtic.

Table 4

Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient ($\rho = +.4626$, N = 30) from the Second Experiment of Egon S. Pearson's Study: 1929

r	Obtained Frequency	Theoretical Frequency
+.82	1	0.3 (+.82 and above)
+.78	5	1.3
+.74	5	4.0
+.70	5	8.6
+.66	15	14.8
+.62	19	21.4
+.58	28	27.2
+.54	33	31.0
+.50	37	32.4
+.46	30	31.5
+.42	31	28.7
+.38	18	24.8
+.34	20	20.4
+.30	15	16.0
+.26	17	12.1
+.22	8	8.8
+.18	6	6.1
+.14	1	4.1
+.10	3	2.7
+.06	1	1.6
+.02	2	(+.02 and below) 2.2
Total	300	300.0

Hilda Frost Dunlap's Study

Hilda Frost Dunlap[2] conducted an empirical study of the distribution of correlation coefficients of samples drawn from rectangular populations. She shuffled a deck of fifty-two playing cards and then dealt them one at a time assigning ranks consecutively to each of the cards as they were dealt. The cards were then shuffled again and the assigned ranks for each of the cards were recorded. She repeated this procedure until she had obtained recordings of sixty rank distributions. Rank order correlation coefficients were calculated for all possible pairs of the sixty rank distributions; i.e., 1,770. The formula for the rank order correlation coefficient was equivalent to the formula for the product moment correlation coefficient since it was impossible to have had any ties.

---

[2] Hilda Frost Dunlap, "An Empirical Determination of the Distribution of Means, Standard Deviations and Correlation Coefficients Drawn from

Dunlap found that the sampling distribution was symmetrical; however, it was slightly leptokurtic. The mean of the correlation coefficients was -.0012 which was not considered significantly different from its parameter of .0000. The theoretical standard error was considered to be $\sqrt{(1-\rho^2)/N}$, which was equal to .1387. The standard error of the obtained sampling distribution was .1359, which was considered a very close approximation of the above value of .1387.

Dunlap concluded that the sampling distribution of product moment correlation coefficients of sample size fifty-two drawn from a rectangular population with a correlation of zero is symmetrical and slightly leptokurtic. She also concluded that the standard error was adequately represented by $\sqrt{(1-\rho^2)/N}$.

Rectangular Populations," *The Annals of Mathematical Statistics*, II, No. 1 (1931), 66-81.

George A. Baker's Study

George A. Baker[3] conducted an experiment on the effects of a skewed population upon its sampling distributions of product moment correlation coefficients when its parameter is equal to zero. A population of size ten thousand was established having two identical positively skewed marginal distributions. Proportional frequencies for the marginal distributions are listed in Table 5.

Fifty samples of size forty were randomly drawn by using a table of random numbers. The obtained correlation coefficients are listed in Table 6. Assuming the standard error of a sampling distribution of product moment correlation coefficients was equal to $\sqrt{(1 - \rho^2)/(N - 1)}$, Baker grouped the obtained correlation coefficients into intervals of one-half of a standard error. The obtained grouped frequency distribution of obtained correlation coefficients is given in Table 7. The distribution of obtained correlation coefficients is skewed.

Baker concluded that the results of the experiment suggested that the sampling distribution of the product moment correlation coefficients from a population such as used in this experiment and for samples of around size forty is skewed. He concluded that correlation coefficients three and one-half to four standard errors from the mean may be nonsignificant. He stated that in dealing with correlation coefficients from a population with very extreme variates, the individual correlation coefficients must be much larger than those expected from a bivariate normal population before they can be considered significant.

---

[3] George A. Baker, "The Significance of the Product-Moment Coefficient of Correlation with Special Reference to the Character of the Marginal Distributions," Journal of the American Statistical Association, XXV, No. 172 (1930), 387-396.

Table 5

Proportional Frequencies for the Marginal Distributions for the Skewed Population Used in George A. Baker's Study

X	Proportional Frequency
22	.08
21	.01
20	.01
19	.01
18	.01
17	.01
16	.01
15	.01
14	.01
13	.01
12	.01
11	.01
10	.01
9	.02
8	.02
7	.02
6	.04
5	.04
4	.06
3	.26
2	.24
1	.10
Total	1.00

Table 6

Obtained Correlation Coefficients

from George A. Baker's Study

r	r
+.671	-.023
+.388	-.024
+.373	-.026
+.364	-.033
+.347	-.046
+.334	-.064
+.303	-.066
+.296	-.072
+.253	-.075
+.245	-.078
+.223	-.095
+.221	-.095
+.214	-.097
+.183	-.097
+.183	-.097
+.183	-.097
+.173	-.098
+.137	-.125
+.112	-.130
+.101	-.139
+.025	-.150
+.0004	-.195
-.007	-.211
-.014	-.217
-.017	-.248
-.022	-.315

Table 7

Grouped Sampling Distribution of Product Moment Correlation Coefficients from George A. Baker's Study

Standard Error	Grouped Frequency
+4.0 to +4.5	1
+3.5 to +4.0	0
+3.0 to +3.5	0
+2.5 to +3.0	0
+2.0 to +2.5	5
+1.5 to +2.0	4
+1.0 to +1.5	6
+ .5 to +1.0	3
0 to + .5	2
- .5 to 0.0	14
-1.0 to - .5	10
-1.5 to -1.0	3
-2.0 to -1.5	2
Total	50

Egon S. Pearson's Study: 1931

Egon S. Pearson[4] explored the effects of nonnormality upon the sampling distributions of the product moment correlation coefficient. He chose four populations having no correlation. The two marginal distributions were identical for each of the populations. The marginal distributions of population one had a coefficient of skewness of .00 and a coefficient of kurtosis of 2.50. The marginal distributions of population two had a coefficient of skewness of .00 and a coefficient of kurtosis of 7.05. The marginal distributions of population three had a coefficient of skewness of .49 and a coefficient of kurtosis of 3.72. And the marginal distributions of population four had a coefficient of skewness of .99 and a coefficient of kurtosis of 3.83.

---

[4] Egon S. Pearson, "The Test of Significance for the Correlation Coefficient," *Journal of the American Statistical Association*, XXVI, No. 174 (1931), 128-134.

Two hundred and fifty random samples of size ten were drawn from each of the populations. Two hundred and fifty random samples of size twenty were drawn from populations one, two, and four and 395 random samples of size twenty were drawn from population three. The obtained sampling distributions together with their expected sampling distributions are listed in tables 8 and 9 along with the probability of occurrence of each obtained distribution as a random sample from its theoretical distribution. Chi-square tests were used to test for goodness of fit. The standard errors for the sampling distributions are listed in Table 10.

Pearson found fairly close agreement of the obtained sampling distributions to the theoretical expected sampling distributions. For sample size ten there were too few high values from the platykurtic population (population one) and too many from the leptokurtic population (population two).

Table 8

Obtained and Theoretical Sampling Distributions
of Product Moment Correlation Coefficients
Together with Tests of Goodness of Fit
($\rho = 0$, N = 10)
from Egon S. Pearson's Study: 1931

r	Expected Frequency	Obtained Frequency Population 1	2	3	4
> .80	1.3	1	1	4	2
.75-.80	1.8	0	4	1	0
.70-.75	2.9	1	3	0	2
.65-.70	4.5	5	10	5	4
.60-.65	6.1	5	2	2	6
.55-.60	8.3	12	9	9	6
.50-.55	10.4	5	13	6	9
.45-.50	12.7	12	10	12	10
.40-.45	15.0	18	18	24	15
.35-.40	17.4	16	21	18	25
.30-.35	19.5	27	13	15	16
.25-.30	21.6	29	24	17	22
.20-.25	23.4	19	21	26	25
.15-.20	24.9	24	20	28	23
.10-.15	26.1	19	29	24	30
.05-.10	26.8	30	18	29	22
.00-.05	27.3	27	34	30	33
Total	250.0	250	250	250	250
Goodness of Fit	n_1	13	13	13	13
	P	.31	.14	.49	.74

Theoretically, there should be 10.5 correlation coefficients of absolute value greater than or equal to .65 in 250 random samples of

Table 9

Obtained and Theoretical Sampling Distributions of Product Moment Correlation Coefficients Together with Tests of Goodness of Fit

($\rho = 0$, N = 20)

from Egon S. Pearson's Study: 1931

r	Expected Frequency	Obtained Frequency Populations				Expected Frequency
		1	2	3	4	
>.60	1.3	1	2	1	3	2.1
.57-.60	.9	2	0	0	1	1.4
.54-.57	1.3	3	1	0	0	2.1
.51-.54	1.9	5	3	3	2	3.0
.48-.51	2.6	2	3	3	5	4.2
.45-.48	3.6	3	4	5	8	5.6
.42-.45	4.7	6	12	4	10	7.4
.39-.42	6.0	4	3	6	10	9.4
.36-.39	7.5	10	10	11	13	11.8
.33-.36	9.[illegible]	8	9	14	10	14.4
.30-.33	10.8	14	11	7	16	17.1
.27-.30	12.7	12	11	15	16	20.1
.24-.27	14.7	12	12	14	19	23.1
.21-.24	16.5	18	14	15	27	26.1
.18-.21	18.3	16	18	17	38	29.0
.15-.18	20.1	14	20	19	34	31.7
.12-.15	21.6	23	19	14	28	34.2
.09-.12	22.9	22	25	24	41	36.1
.06-.09	23.9	29	23	23	37	37.8
.03-.06	24.6	24	18	34	38	38.9
.00-.03	25.0	22	32	21	39	39.5
Total	250.0	250	250	250	395	395.0
Goodness of Fit	n_1	15	15	15	17	
	P	.85	.55	.83	.83	

Table 10

Obtained and Theoretical Standard Errors of the Product Moment Correlation Coefficient

($\rho = 0$, N = 10, N = 20)

from Egon S. Pearson's Study: 1931

Population	N = 10	N = 20
1	.3234	.2413
2	.3487	.2405
3	.3238	.2289
4	.3181	.2308
Theoretical	.3333	.2294

size ten. The obtained frequencies for the four populations were 7 from population one, 18 from population two, 10 from population three, and 8 from population four.

Theoretically, there should be 11.6 correlation coefficients of absolute value greater than or equal to .45 in 250 samples of size twenty. The obtained frequencies for the four populations were 16 from population one, 13 from population two, 12 from population three, and 12 from population four. The frequency for population three was reduced to conform to a sampling distribution of 250 samples instead of 395 samples.

Pearson concluded that although the sampling distribution frequencies were too small to draw conclusive evidence, the results of the study did indicate that the product moment correlation coefficient is remarkably insensitive to changes in the form of the population.

Egon S. Pearson's Study: 1932

Egon S. Pearson[5] conducted an extension of his earlier study[6] on the effects of nonnormality upon the sampling distributions of the product moment correlation coefficient. The two marginal distributions were identical for each of the populations in his earlier study. The two marginal distributions were not identical for each of the populations in this study. He constructed three populations using the univariate distributions he had used in the earlier study. Population A had marginal distributions from populations one (coefficient of skewness equal to .00 and coefficient of kurtosis equal to 2.50) and two (coefficient of skewness equal to .00 and coefficient of kurtosis equal to 7.05). Population B had marginal distributions from populations one and four (coefficient of skewness equal to .99 and coefficient of kurtosis equal to 3.83). And population C had marginal distributions from populations two and four.

Five hundred random samples of size ten and five hundred random samples of size twenty were drawn from each of the populations. The obtained and expected sampling distributions are listed in tables 11 and 12 together with the probability of occurrence of each obtained distribution as a random sample from its theoretical distribution. Chi-square

---

[5] Egon S. Pearson, "The Test of Significance for the Correlation Coefficient: Some Further Results," *Journal of the American Statistical Association*, XXVII, No. 180 (1932), 424-426.

[6] Egon S. Pearson, "The Test of Significance for the Correlation Coefficient," *Journal of the American Statistical Association*, XXVI, No. 174 (1931), 128-134.

Table 11

Obtained and Theoretical Sampling Distributions

of Product Moment Correlation Coefficients

Together with Tests of Goodness of Fit

($\rho = 0$, $N = 10$)

from Egon S. Pearson's Study: 1932

r	Expected Frequency	Obtained Frequency Population A	B	C
> .80	2.7	3	1	2
.75-.80	3.6	1	2	3
.70-.75	5.8	3	4	8
.65-.70	8.9	11	5	7
.60-.65	12.3	11	10	8
.55-.60	16.5	20	17	18
.50-.55	20.8	28	14	20
.45-.50	25.3	24	27	19
.40-.45	30.1	21	35	26
.35-.40	34.7	36	29	34
.30-.35	39.1	44	44	51
.25-.30	43.2	37	47	34
.20-.25	46.8	38	58	59
.15-.20	49.8	57	57	47
.10-.15	52.1	54	45	54
.05-.10	53.7	55	58	51
.00-.05	54.6	57	47	59
Total	500.0	500	500	500
Goodness of Fit	n	15	15	15
	P	.58	.41	.47

Table 12

Obtained and Theoretical Sampling Distributions of Product Moment Correlation Coefficients Together with Tests of Goodness of Fit

($\rho = 0$, N = 20)

from Egon S. Pearson's Study: 1932

r	Expected Frequency	Obtained Frequency Population A	B	C
> .60	2.6	1	2	4
.57-.60	1.8	1	4	2
.54-.57	2.6	4	0	3
.51-.54	3.8	2	5	1
.48-.51	5.3	4	5	4
.45-.48	7.1	10	5	6
.42-.45	9.4	11	6	7
.39-.42	12.0	10	10	11
.36-.39	14.9	15	15	20
.33-.36	18.2	19	20	15
.30-.33	21.7	21	28	21
.27-.30	25.4	29	26	23
.24-.27	29.3	29	34	25
.21-.24	33.0	29	32	43
.18-.21	36.7	44	34	42
.15-.18	40.1	38	32	34
.12-.15	43.3	37	54	57
.09-.12	45.7	42	41	38
.06-.09	47.9	60	46	50
.03-.06	49.2	47	49	51
.00-.03	50.0	47	52	43
Total	500.0	500	500	500
Goodness of Fit	n	18	18	18
	P	.91	.89	.50

tests were used to test for goodness of fit. The standard errors of the sampling distributions are listed in Table 13.

Table 13

Obtained and Theoretical Standard Errors of the

Product Moment Correlation Coefficient

($\rho = 0$, N = 10, N = 20)

from Egon S. Pearson's Study: 1932

Population	N = 10	N = 20
A	.3299	.2276
B	.3161	.2269
C	.3235	.2260
Theoretical	.3333	.2294

Theoretically, there should be 21.0 samples with correlation coefficients of absolute value greater than or equal to .65 in five hundred random samples of size ten. The obtained frequencies were 18 for population A, 12 for population B, and 20 for population C.

Theoretically, there should be 23.2 samples with correlation coefficients of absolute value greater than or equal to .45 in five hundred samples of size twenty. The obtained frequencies were 22 for population A, 21 for population B, and 20 for population C.

In all cases there were fewer obtained correlation coefficients than expected correlation coefficients. The differences were, however, very small. The largest difference occurred for sample size ten for population B. One of the marginal distributions for population B was platykurtic and the other was very skewed. This discrepancy did not reoccur for sample size twenty.

Pearson concluded that this study again emphasized the insensitiveness of the sampling distributions of the product moment correlation coefficient to changes in population form.

Leone Chesire, Elena Oldis, and Egon S. Pearson's Study

Leone Chesire, Elena Oldis, and Egon S. Pearson[7] conducted a study on the effect of nonnormality upon the sampling distributions of the product moment correlation coefficient. They constructed two populations. They assumed that if two dice were thrown, the sum of the two dice would give an X value for the first population. If one of the thrown dice was chosen randomly and thrown again, the sum of the two dice on the table would give a Y_1 value for the first population. A bivariate frequency table of the expected X and Y_1 values was constructed for a total frequency of 216. The first population is listed in Table 14. The marginal distributions of X and Y_1 were the same and were triangular. The regression was linear. The correlation of the population was .500. The arrays were not homoscedastic.

They further assumed that if two dice were thrown, the sum of the two dice would give an X value for the second population. Then if the

---

[7] Leone Chesire, Elena Oldis, and Egon S. Pearson, "Further Experiments on the Sampling Distribution of the Correlation Coefficient," *Journal of the American Statistical Association*, XXVII, No. 178 (1932), 121-126.

Table 14

First Nonnormal Population Having a Correlation of .50 from Leone Chesire, Elena Oldis, and Egon S. Pearson's Study

Y_1 Value	X Value 2	3	4	5	6	7	8	9	10	11	12	Total
12						1	1	1	1	1	1	6
11					1	2	2	2	2	2	1	12
10				1	2	3	3	3	3	2	1	18
9			1	2	3	4	4	4	3	2	1	24
8		1	2	3	4	5	5	4	3	2	1	30
7	1	2	3	4	5	6	5	4	3	2	1	36
6	1	2	3	4	5	5	4	3	2	1		30
5	1	2	3	4	4	4	3	2	1			24
4	1	2	3	3	3	3	2	1				18
3	1	2	2	2	2	2	1					12
2	1	1	1	1	1	1						6
Total	6	12	18	24	30	36	30	24	18	12	6	216

die which was lower in value or equal in value to the other die was thrown again, the two dice on the table would give a Y_2 value for the second population. A bivariate frequency table of the expected X and Y_2 values was constructed for a total frequency of 126. The second population is listed in Table 15. The distribution of X for both populations was the same. The Y_2 distribution was asymmetrical. The regression was not

Table 15

Second Nonnormal Population Having a Correlation of .5462 from

Leone Chesire, Elena Oldis, and Egon S. Pearson's Study

X_2 Value	X Value											Total
	2	3	4	5	6	7	8	9	10	11	12	
12						2	2	2	2	2	1	11
11					2	4	4	4	3	2	1	20
10				2	4	6	5	4	3	2	1	27
9			2	4	5	6	5	4	3	2	1	32
8		2	3	4	5	6	5	4	3	2	1	35
7	1	2	3	4	5	6	5	4	3	2	1	36
6	1	2	3	4	5	4	3	2	1			25
5	1	2	3	4	3	2	1					16
4	1	2	3	2	1							9
3	1	2	1									4
2	1											1
Total	6	12	18	24	30	36	30	24	18	12	6	216

linear. The correlation of the second population was .5462. The array distributions were varied.

The sampling procedure employed was the same for both populations. Each of the 216 bivariates was placed on a card. The cards were shuffled. Five cards were drawn to comprise a sample of size five. The first card was taken from the deck about one-half of the distance from the top, the second card about one-sixth of the distance from the top, the third card about one-third of the distance from the top, the fourth card about two-thirds of the distance from the top, and the fifth card about five-sixths of the distance from the top. The five cards were then replaced near the center of the deck, the cards reshuffled, and the second sample of five cards was drawn. The samples of size five were combined by pairs consecutively to form five hundred samples of size ten. The samples of size ten were combined by pairs consecutively to form 250 samples of size twenty. Thus, the sampling distributions of sample sizes 5, 10, and 20 were not independent of one another.

The obtained and theoretical sampling distributions of the product moment correlation coefficients are given for sample size five in Table 16, for sample size ten in Table 17, and for sample size twenty in Table 18. The obtained and theoretical values for the mean correlation coefficients of the sampling distributions, the obtained and theoretical standard errors of the sampling distributions, and the probability of occurrence of each obtained distribution as a random sample from its theoretical distribution are listed in Table 19. Chi-square tests were used to test for goodness of fit.

Chesire, Oldis, and Pearson concluded that the observed and theoretical values of the means of the sampling distributions were in quite close agreement, but that the standard errors of the obtained sampling distributions were less than the expected theoretical values. They observed that the sampling distributions for the second population approximated normality more closely than that for the first population. They also concluded that the agreement between the obtained distributions

Table 16

Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient for Sample Size Five from Leone Chesire, Elena Oldis, and Egon S. Pearson's Study

r	First Population		Second Population	
	Obtained	Theoretical	Obtained	Theoretical
+.95	85	98.9	99	120.0
+.85	117	131.8	154	148.9
+.75	122	124.2	124	132.5
+.65	113	109.0	114	111.2
+.55	98	93.2	93	91.6
+.45	101	78.8	84	75.1
+.35	85	66.3	68	61.7
+.25	48	55.7	67	50.7
+.15	52	46.7	48	41.8
+.05	12	39.2	34	34.5
-.05	35	32.8	25	28.5
-.15	24	27.4	16	23.5
-.25	19	22.9	17	19.4
-.35	15	19.0	17	16.0
-.45	12	15.6	12	13.0
-.55	16	12.7	13	10.5
-.65	6	10.1	4	8.3
-.75	4	7.7	4	6.3
-.85	2	5.4	6	4.4
-.95	4	2.6	1	2.1
Total	1,000	1,000.0	1,000	1,000.0

and the theoretical distributions increased as the sample size increased. They concluded that there were real differences between sampling distributions of the correlation coefficients from these two populations and the theoretical ones from normal populations. They stated that whether or not the difference is significant depends upon the degree of accuracy desired.

Table 17

Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient for Sample Size Ten from Leone Chesire, Elena Oldis, and Egon S. Pearson's Study

r	First Population		Second Population	
	Obtained	Theoretical	Obtained	Theoretical
+.975	1	.7	0	.9
+.925	6	4.6	6	6.7
+.875	14	12.9	20	17.9
+.825	18	22.4	29	29.5
+.775	20	30.5	31	38.0
+.725	40	36.8	45	43.3
+.675	30	40.1	43	44.9
+.625	42	41.1	51	42.9
+.575	50	40.2	42	41.1
+.525	37	38.0	39	37.4
+.475	39	34.9	37	33.1
+.425	38	31.5	22	28.9
+.375	36	27.9	23	24.8
+.325	28	24.3	18	21.0
+.275	21	20.9	25	17.5
+.225	16	17.7	22	14.6
+.175	13	14.9	13	12.0
+.125	15	12.4	10	9.7
+.075	9	10.2	5	7.8
+.025	13	8.4	9	6.3
-.025	2	6.8	1	5.0
-.075	0	5.4	1	3.9
-.125	8	4.3	2	3.0
-.175	1	3.4	3	2.4
-.225	1	2.6	0	1.8
-.275	1	2.0	1	1.3
-.325	1	1.5	0	1.0
-.350 & less	0	3.6	2	2.3
Total	500	500.0	500	500.0

Table 18

Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient for Sample Size Twenty from Leone Cheshire, Elena Oldis, and Egon S. Pearson's Study

r	First Population		Second Population	
	Obtained	Theoretical	Obtained	Theoretical
+.975	0	.0	0	.0
+.925	1	.1	0	.1
+.875	1	.6	1	1.0
+.825	3	2.9	4	5.2
+.775	5	7.8	10	12.7
+.725	12	14.5	23	21.0
+.675	24	21.2	30	27.7
+.625	24	26.3	37	31.0
+.575	24	28.8	26	30.8
+.525	33	28.7	29	28.0
+.475	31	26.4	24	23.9
+.425	21	22.9	24	19.3
+.375	15	18.8	14	14.8
+.325	19	14.8	15	11.1
+.275	17	11.2	3	7.9
+.225	11	8.2	5	5.5
+.175	1	5.8	2	3.7
+.125	3	4.0	2	2.5
+.075	2	2.7	1	1.6
+.025	1	1.7	0	.9
-.025	1	1.1	0	.6
-.075	1	.7	0	.3
<-1.00	0	.8	0	.4
Total	250	250.0	250	250.0

Table 19

Obtained and Theoretical Mean Product Moment Correlation Coefficients, Standard Errors of the Product Moment Correlation Coefficient, and Results of Tests of Goodness of Fit of Obtained and Theoretical Sampling Distributions of Product Moment Correlation Coefficients from Leone Chesire, Elena Oldis, and Egon S. Pearson's Study

		Sample Size Five		Sample Size Ten		Sample Size Twenty	
		First Pop.	Second Pop.	First Pop.	Second Pop.	First Pop.	Second Pop.
Mean r	Obtained	.4542	.4924	.4871	.5289	.4896	.5460
	Theoretical	.4517	.4958	.4787	.5242	.4900	.5359
Std. Error	Obtained	.3961	.3924	.2371	.2383	.1701	.1503
	Theoretical	.4239	.4076	.2671	.2534	.1780	.1677
Goodness of Fit	n	18	18	23	22	15	14
	P	.122	.374	.096	.492	.406	.633

Paul R. Rider's Study

Paul R. Rider[8] conducted an investigation of the sampling distributions of the product moment correlation coefficient in small samples randomly drawn from nonnormal populations. He constructed three populations. One population was normal with seven intervals in each marginal distribution. The correlation of the bivariate normal distribution was .83. The coefficient of skewness for the marginal distributions was zero and the coefficient of kurtosis for the marginal distributions was 2.9390. A second population was rectangular. It consisted of a ten by ten scatterplot with equal frequencies in each cell. The correlation of the rectangular population was zero. The coefficient of skewness for the marginal distributions was zero and the coefficient of kurtosis for the marginal distributions was 1.7758. A third population was triangular. It consisted of a ten by ten scatterplot. The frequencies of all of the cells on one side of a principal diagonal were zero and the frequencies of the remaining cells were all equal to a constant value other than zero. The correlation of the traingular population was one-half. The coefficient of skewness of the marginal distributions was .326 and the coefficient of kurtosis of the marginal distributions was 2.36.

---

[8] Paul R. Rider, "On the Distribution of the Correlation Coefficient in Small Samples," Biometrika, XXIV, Parts 3 and 4 (1932), 382-403.

The sampling was done by using a table of random numbers. One thousand samples of size five were drawn from each of the three populations. By combining the samples of size five, five hundred samples of size ten were obtained for the triangular and the normal populations. The obtained sampling distributions are listed in Table 20. The obtained mean correlation coefficients and the obtained standard error of the correlation coefficient together with the theoretical normal theory values are listed in Table 21.

Table 20

Obtained and Theoretical Sampling Distributions

from Paul R. Rider's Study

	Rectangular		Triangular				Normal	
	N = 5		N = 5		N = 10		N = 5	N=10
r	Ob	Th	Ob	Th	Ob	Th	Ob	Ob
+ .95 to +1.00	9	6.2	48	98.9	3	0.7	235	34
+ .90 to + .95	10	11.9	59		9	4.6	175-1/2	97
+ .85 to + .90	17	15.4	49	131.8	11	12.9	146-1/2	92
+ .80 to + .85	20	18.0	62		17	22.4	83	86
+ .75 to + .80	19	20.1	64	124.2	28	30.5	98-1/2	46
+ .70 to + .75	22	21.9	59		27	36.8	33-1/2	56
+ .65 to + .70	35	23.4	64	109.0	40	40.1	36	23
+ .60 to + .65	21	24.9	47		46	41.1	65	21
+ .55 to + .60	33	26.0	55	93.2	54	40.2	12	15
+ .50 to + .55	23	27.1	46		31	38.0	27	15
+ .45 to + .50	23	28.0	53	78.8	36	34.9	14	4
+ .40 to + .45	27	28.7	50		36	31.5	21	3
+ .35 to + .40	33	29.6	38	66.3	38	27.9	2	3
+ .30 to + .35	30	30.1	25		19	24.3	14	1
+ .25 to + .30	27	30.5	40	55.7	22	20.9	4	0
+ .20 to + .25	26	31.1	27		22	17.7	4	2
+ .15 to + .20	33	31.3	31	46.7	15	14.9	8	0
+ .10 to + .15	32	31.6	14		22	12.4	4	0
+ .05 to + .10	33	31.7	20	39.2	6	10.2	1	0
.00 to + .05	24	31.8	16-1/2		2-1/2	8.4	4	0
− .05 to .00	33	31.8	17-1/2	32.8	4-1/2	6.8	4	0
− .10 to − .05	29	31.7	15		2	5.4	0	1
− .15 to − .10	32	31.6	4	27.4	1	4.3	0	1
− .20 to − .15	28	31.3	9		0	3.4	4	0
− .25 to − .20	29	31.1	9-1/2	22.9	3	2.6	4	0
− .30 to − .25	28	30.5	11-1/2		3	2.0	2	0
− .35 to − .30	24	30.1	11	19.0	0	1.5	0	0
− .40 to − .35	28	29.6	5		0		0	0
− .45 to − .40	27	28.7	8	15.6	0		4	0
− .50 to − .45	41	28.0	6		0		0	0
− .55 to − .50	24	27.1	5	12.7	0		0	0
− .60 to − .55	28	26.0	6		0		0	0
− .65 to − .60	29	24.9	4	10.1	1	3.6	0	0
− .70 to − .65	21	23.4	2		0		0	0
− .75 to − .70	21	21.9	5-1/2	7.7	0		0	0
− .80 to − .75	16	20.1	6-1/2		0		0	0

Table 20 (continued)

r	Rectangular		Triangular				Normal	
	N = 5		N = 5		N = 10		N = 5	N=10
	Ob	Th	Ob	Th	Ob	Th	Ob	Ob
- .85 to - .80	17	18.0	3	5.4	0		0	0
- .90 to - .85	27	15.4	1		0		0	0
- .95 to - .90	9	11.9	1	2.6	1		0	0
-1.00 to - .95	12	6.7	2		0		0	0
Total	1000	999.6	1000	1000.0	500		1000	500

Table 21

Obtained and Theoretical Mean Correlation Coefficients

and Standard Errors of the Correlation Coefficient

from Paul R. Rider's Study

Population	Sample Size	Obtained or Theoretical	Mean r	Standard Error
Rectangular	5	Obtained	-.0031	.5153
Rectangular	5	Theoretical	+.0000	.5000
Traingular	5	Obtained	+.4653	.4127
Triangular	5	Theoretical	+.4517	.4239
Triangular	10	Obtained	+.4910	.2483
Triangular	10	Theoretical	+.4787	.2671
Normal	5	Obtained	+.7852	.2306
Normal	5	Theoretical	+.7873	.2461
Normal	10	Obtained	+.8012	.1433
Normal	10	Theoretical	+.8135	.1288

Rider applied tests of goodness of fit to some of the obtained sampling distributions and the expected normal theory distributions. A χ^2 of 39.5168, $n = 39$, was obtained for sample size five of the rectangular population giving $p = .448$. A χ^2 of 25.51, $n = 18$, was obtained for sample size five of the triangular population giving $p = .112$. A χ^2 of 46.89, $n = 22$, was obtained for sample size ten of the triangular population giving $p = .00153$. This was the only extremely bad fit. A χ^2 of 23.47, $n = 12$, was obtained for sample size ten of the normal population giving $p = .024$. This fit was not too good, but the main discrepancies were somewhat irregularly scattered throughout the distribution. Rider used Fisher's transformations for the sampling distributions in testing the goodness of fit in this last case.

Rider concluded that the obtained sampling distributions of the product moment correlation coefficient of random samples from nonnormal populations were not essentially different from the theoretical normal theory sampling distributions.

Conclusions

The general conclusion drawn by the researchers of the above studies is that the sampling distributions of the product moment correlation coefficient are fairly insensitive to the effects of nonnormality in regards to the shape of the population. However, the number of correlation coefficients in the obtained sampling distributions are too few to make conclusive statements as to the effects of nonnormality upon the sampling distributions of the product moment correlation coefficient.

CHAPTER III

PROCEDURE

Dee Wiswall Norton[1] conducted an empirical investigation of the effects of nonnormality and heterogeneity of variance on the F-distribution using a procedure similar to the one used in this study. The process of establishing large populations on IBM cards and of selecting random samples was patterned after Norton's recommendations.

The general procedure in the present study was to establish large finite bivariate populations. The bivariates were punched on IBM cards. The IBM cards were placed in random order by ordering random digits which had also been punched on the IBM cards. Samples of specified sizes were counted off, product moment correlation coefficients were calculated for the samples, and sampling distributions were established for the correlation coefficients of the numerous samples. The IBM 650 Data Processing System was used for sampling and calculating.

Ten finite bivariate populations of size ten thousand were established. Five distributions were established having approximately no correlation and five having substantial correlation. Two of the ten populations were control groups; i.e., they were bivariate normal, and eight of the populations were experimental groups; i.e., they were bivariate nonnormal.

---

[1] Dee Wiswall Norton, "An Empirical Investigation of the Effects of Nonnormality and Heterogeneity upon the F-test of Analysis of Variance" (Unpublished Ph.D. dissertation, State University of Iowa, 1952).

The procedure used in establishing the bivariate frequencies was to establish the basic univariate distribution along the abscissa. The variates at each interval along the abscissa were then assigned their ordinate values. The variance of the ordinate values for the variates at each interval of the abscissa was kept approximately equal to the variance of the entire distribution along the abscissa in establishing the five populations having approximately no correlation. The variance of the ordinate values for the variates at each interval of the abscissa was kept approximately equal and at a calculated fraction of the variance of the entire distribution along the abscissa in establishing the five populations having substantial correlation. By following the above procedure each population would have the property of homogeneity of variance from column to column within limitations due to the populations being finite.

The shape of each columnar array was kept approximately similar to the shape of the entire distribution along the abscissa. Thus, each distribution would be homoclitic among the columnar arrays and also homokurtic among the columnar arrays within limitations due to the populations being finite.

The columnar arrays for the populations were placed in order to make the correlation of the populations approximately linear.

Five basic univariate distributions of size ten thousand were established. They were normal, rectangular, leptokurtic, slightly skewed, and markedly skewed. The means, variances, coefficients of skewness, and coefficients of kurtosis for these five basic univariate distributions are listed in Table 22. Each of the above univariate distributions was used

Table 22

Characteristics of Univariate Distributions Used

X	Univariate Distribution	Mean	Variance	Coefficient of Skewness	Coefficient of Kurtosis
X_1	Normal	9.50	6.66	.00	2.99
X_2	Rectangular	9.50	33.25	.00	1.79
X_3	Leptokurtic	9.50	1.77	.00	6.47
X_4	Slightly Skewed	5.82	6.33	+.69	3.68
X_5	Markedly Skewed	3.91	9.61	+.99	3.82

in establishing a bivariate population having approximately no correlation and in a bivariate population having substantial correlation.

The formula for computing the coefficients of skewness was $(1/N)\Sigma((X - \bar{X})/\sigma_x)^3$. A perfectly symmetrical distribution will have a coefficient of skewness equal to .00, a negatively skewed distribution will have a negative value, and a positively skewed distribution will have a positive value.

The formula for computing the coefficients of kurtosis was $(1/N)\Sigma((X - \bar{X})/\sigma_x)^4$. A mesokurtic distribution will have a coefficient of kurtosis equal to 3.00, a platykurtic distribution will have a coefficient less than 3.00, and a leptokurtic distribution will have a coefficient greater than 3.00.

Tables of the ten bivariate populations of size ten thousand are listed in Appendix A. The correlation coefficients of the ten populations are listed in Table 23.

Table 23

Correlation Coefficients of the Established Populations

Population	ρ	Distribution along the Abascissa
Normal	.000	X_1
Rectangular	.000	X_2
Leptokurtic	.000	X_3
Slightly Skewed	-.005	X_4
Markedly Skewed	-.005	X_5
Normal	+.832	X_1
Rectangular	+.834	X_2
Leptokurtic	+.813	X_3
Slightly Skewed	+.828	X_4
Markedly Skewed	+.832	X_5

The bivariates and cross products for all ten of the populations were punched on the deck of ten thousand detail cards. Lower cumulative frequencies by cells were determined for each of the ten populations. The frequencies in the cells were cumulated by starting with the cell in the lower left hand corner, proceeding up the column cell by cell, returning to the lowest cell of the adjacent column to the right, proceeding up this column cell by cell, etc. For each lower cumulative frequency a master card was punched with the value of the lower cumulative frequency and the values of the variates and their cross product. One set of master cards was prepared for all ten populations. Therefore, every time a unique lower cumulative frequency was obtained in any one of the ten populations, a new master card was prepared. The new master card contained the lower cumulative frequency and the values of the variates and their cross product plus the values of the variates and their cross products for the other populations at that cumulative frequency.

Norton[2] collated the master cards using the lower cumulative frequency values with the deck of ten thousand detail cards using serial numbers 0000 to 9999. By following this procedure the card populations punched on the deck of ten thousand detail cards would be identical to the originally established populations. However, Norton recommended an alternative procedure to be followed in preference to the procedure which he had used in his study. He recommended punching four columns of random digits on the deck of ten thousand detail cards and then ordering the four-digit random numbers. The deck of ten thousand detail cards would then be similar but not necessarily identical to the deck of ten thousand cards with serial numbers 0000 to 9999 punched on them. Under Norton's recommended procedure, bivariate distributions of ten thousand would in effect be considered samples of size ten thousand drawn from infinite populations of which the originally established distributions of size ten thousand would be representative. In the procedure used by Norton the probability of a value whose relative frequency is .0001 occurring in more than one sample in a single ordering of the ten thousand detail cards is equal to zero. However, in the recommended procedure, the probability of its occurring a second time is not equal to zero.

---

[2] Dee Wiswall Norton, "An Empirical Investigation of the Effects of Nonnormality and Heterogeneity upon the F-test of Analysis of Variance" (Unpublished Ph.D. dissertation, State University of Iowa, 1952), p. 147.

The alternative procedure recommended by Norton was used in this study. The master cards using the lower cumulative frequencies were collated with the ten thousand detail cards using a set of ordered four-digit random numbers. The values of the bivariates and their cross products were then gang punched from the master cards onto the detail cards. The columns assigned to the fields on the master cards and the detail cards are listed in Table 24.

The random digits needed for the detail cards were reproduced from IBM cards with random digits already punched on them. These cards were obtained from the Rand Corporation, Santa Monica, California.[3]

The deck of ten thousand detail cards was placed in random order by ordering a set of four-digit random numbers. The samples were obtained by repeatedly counting off the number of cards needed for the samples until the deck of ten thousand detail cards was exhausted. The deck of ten thousand cards was then placed in another unique random order by ordering another set of four-digit random numbers. Additional samples were then obtained by again repeatedly counting off the number of cards needed for the samples. The deck of ten thousand detail cards was continually

---

3. Rand Corporation, A Million Random Digits with 100,000 Normal Deviates (Glencoe, Illinois: Free Press, 1955).

Table 24

Fields on Master Cards and Detail Cards

Field	Columns	
	Master Cards	Detail Cards
X_1	16-17	16-17
X_2	18-19	18-19
X_3	20-21	20-21
X_4	22-23	22-23
X_5	24-25	24-25
X_6	26-27	26-27
X_1X_6	28-30	28-30
X_7	31-32	31-32
X_2X_7	33-35	33-35
X_8	36-37	36-37
X_3X_8	38-40	38-40
X_9	41-42	41-42
X_4X_9	43-45	43-45
X_{10}	46-47	46-47
X_5X_{10}	48-50	48-50
X_{11}	51-52	51-52
X_1X_{11}	53-55	53-55
X_{12}	56-57	56-57
X_2X_{12}	58-60	58-60
X_{13}	61-62	61-62
X_3X_{13}	63-65	63-65
X_{14}	66-67	66-67
X_4X_{14}	68-70	68-70
X_{15}	71-72	71-72
X_5X_{15}	73-75	73-75
Random Digits	-----	1-15, 76-80
Lower Cumulative Frequency	2-5	------

rerandomized and samples counted off until the desired number of samples had been obtained.

The number of samples required to determine a sampling distribution representative of the sampling distribution of an infinite number of samples cannot be determined. However, there is usually more concern about the tails of the sampling distribution than about the center of the distribution. The .01 and .05 levels of significance are very commonly used in educational and psychological research. If it is assumed that the probability of making an alpha error in repeated sets of samples is normally distributed and its standard error is $\sqrt{pq/N}$ where N is equal to the number of samples, the size of N needed to obtain a certain degree of precision can be determined. If the degree of precision desired is such that an alpha error would be expected to occur .015 to .005 of the time in repeated sets of samples ninety-nine per cent of the time when working at the .01 level of significance, the standard error would be equal to .005/2.58. Substituting this value into the above formula and solving for N,

$$\sigma_p = \sqrt{pq/N}$$

$$\sigma_p = .005/2.58$$

$$p = .01$$

$$q = .99$$

$$.005/2.58 = \sqrt{(.01)(.99)/N}$$

$$N = 2{,}636.$$

Thus, at least 2,636 samples were obtained for each sampling distribution in order to obtain this degree of precision.

Sampling distributions of correlation coefficients were obtained for each of the ten populations for sample sizes 15, 30, and 90. The samples of sizes 15, 30, and 90 were drawn concurrently. The first fifteen cards comprised the first sample of size fifteen. The second fifteen cards comprised the second sample of size fifteen. These first thirty cards which comprised the first two samples of size fifteen comprised the first sample of size thirty. Similarly, the cards that comprised the first three samples of size thirty comprised the first sample of size ninety.

In order to obtain at least 2,636 samples of size ninety, the deck of ten thousand detail cards had to be placed in random order and sampled to exhaustion twenty-four times. Four columns of random digits were needed for each ordering of the ten thousand detail cards. Therefore, ninety-six columns of random digits were needed for arranging the detail cards in random order twenty-four times in addition to the four columns of random digits needed for collating the master cards with the detail cards. The bivariates and the cross products for the ten populations required sixty columns. There are eighty columns on an IBM card. This left only twenty columns on the IBM cards for random digits. Therefore, five decks of ten thousand detail cards were prepared. Each deck contained twenty columns of random digits plus sixty columns of data. Each deck was used five times for ordering sets of four columns of random digits. No random digit was used more than once. This was necessary in order to prevent the random orders from being related to one another. The sixty columns of data were identical for the five decks.

In summary, the detail cards were placed in random order by ordering four columns of random digits. Samples were then repeatedly counted off concurrently for sample sizes 15, 30, and 90 and correlation coefficients were calculated for samples drawn from all ten of the populations simultaneously. This procedure was repeated twenty-four times by reordering four columns of random digits each time.

Through use of Norton's recommended procedure of collating the lower cumulative frequencies on the master cards with ordered random digits instead of serial numbers 0000 through 9999, the correlation in the card populations did not remain equal to the correlations in the originally established populations. The correlation coefficients for the originally established populations and for the card populations are given in Table 25.

A major criticism of the sampling procedure is that it involves random sampling without replacement. However, it would not be practical in this study to have replaced a card or even each sample of cards each time it was drawn and to have reordered the deck of ten thousand detail cards before drawing another card.

Another criticism of the procedure is that each obtained sampling distribution is not independent of the others. The ten populations were sampled simultaneously. The bivariates and cross products for the ten populations were punched on one set of master cards. The lower cumulative frequencies for each of the populations were obtained by proceeding from the lower left cell to the upper right cell. Thus, there would be a certain relationship among the values of the bivariates of the ten populations as they were punched on the detail cards.

Also, the obtained sampling distributions for sample sizes 15, 30,

Table 25

Correlations of Established Populations and of Card Populations

Population	Correlation Established Population	 Card Populat
Normal	.000	+.016
Rectangular	.000	-.019
Leptokurtic	.000	-.004
Slightly Skewed	-.005	-.010
Markedly Skewed	-.004	-.012
Normal	+.832	+.835
Rectangular	+.834	+.831
Leptokurtic	+.813	+.809
Slightly Skewed	+.828	+.826
Markedly Skewed	+.832	+.826

and 90 are not independent of each other because of the procedure used in obtaining the various sized samples. If there were, through chance, a large number of low values of product moment correlation coefficients in samples of size fifteen, it would be expected that there should also be a large number of low values in samples of size thirty, and likewise in samples of size ninety.

CHAPTER IV

RESULTS AND CONCLUSIONS

In assessing the effects of nonnormality upon the sampling distribution of the product moment correlation coefficient two comparisons seemed appropriate. First, it seemed desirable to compare the empirically determined sampling distribution with the theoretical sampling distribution based on the assumption that Fishers' z transformation yields normally distributed values. This is the distribution usually used to test hypotheses concerning the population correlation. Second, a comparison of the empirical sampling distribution from nonnormal populations with the empirical distribution from the normal population seemed desirable. In cases where the theoretical and empirical normal distributions differed, this latter comparison would more clearly show the effects of nonnormality.

Deviations of obtained distributions from the theoretical distributions were tested by use of a Kolmogorov-Smirnov goodness-of-fit test[1]. In this test the largest absolute deviation between theoretical and empirical cumulative proportions at the upper limits of distribution intervals were compared with critical values at the .20, .15, .10, .05, and .01 levels of significance. In addition, the critical values of r at the .05 and .01 levels of significance were obtained, and the relative incidence of "significant" correlations in the empirical and theoretical distributions were compared by subjective analysis.

---

[1] Sidney Siegel, Nonparametric Statistics for the Behavioral Sciences (New York: McGraw-Hill Book Company, Inc., 1956), pp. 47-52.

To bring out possible patterns in the differential effects of nonnormality, the comparisons were made separately for various types and degrees of nonnormality, three different sample sizes, and two markedly different population correlations. Since population correlation coefficient appeared to have a more pronounced effect than any of the other variables, the results were compiled separately for each value of the population correlation, rho.

Populations Having Approximately No Correlation

The results of the Kolmogorov-Smirnov tests comparing the obtained sampling distributions with their theoretical distributions are listed in Table 26. The proportions of coefficients in the obtained sampling distributions which were more extreme than the theoretical values of $r_{.005}$, $r_{.025}$, $r_{.975}$ and $r_{.995}$ are listed in tables 27 and 28.

Normal (Control) Population

The obtained sampling distributions for the normal (control) population corresponded very closely to their theoretical distributions. (See Graphs 1, 6, and 11 of Appendix C) The obtained distributions for sample sizes 30 and 90 did not show significant deviation from their theoretical distributions even at the .20 level of significance. For sample size fifteen there was deviation between the obtained and theoretical sampling distributions which could happen between fifteen and twenty percent of the time by chance alone. For a two-tailed test at the .05 level of significance the proportion of correlations in the

Table 26

Results of Kolmogorov-Smirnov Test Comparing Obtained Sampling Distributions from Populations Having Approximately No Correlation with Their Theoretical Sampling Distributions

Population	Population Correlation	N	Number of Samples	Absolute Maximum Deviation
Normal	+.016	15	15,984	.0088*
Rectangular	-.019	15	15,984	.0110***
Leptokurtic	-.004	15	15,984	.0114***
Slightly Skewed	-.010	15	15,984	.0079
Markedly Skewed	-.012	15	15,984	.0091**
Normal	+.016	30	7,992	.0085
Rectangular	-.019	30	7,992	.0078
Leptokurtic	-.004	30	7,992	.0102
Slightly Skewed	-.010	30	7,992	.0107
Markedly Skewed	-.012	30	7,992	.0074
Normal	+.016	90	2,664	.0134
Rectangular	-.019	90	2,664	.0082
Leptokurtic	-.004	90	2,664	.0265***
Slightly Skewed	-.010	90	2,664	.0113
Markedly Skewed	-.012	90	2,664	.0088

*Significant at $\alpha = .20$
**Significant at $\alpha = .15$
***Significant at $\alpha = .05$

separate tails did not deviate from the expected .0250 by more than .0032 even for sample size fifteen, and the proportion in both tails combined did not deviate from the expected .0500 by more than .0027 for any one sample size. At the .01 level of significance the proportion of

Table 27

Proportions of Significant Coefficients in Sampling Distributions from Populations Having Approximately No Correlation ($\alpha = .05$)

Population	Population Correlation	N	$\leq r_{.025}$	$\geq r_{.975}$	$\leq r_{.025}$ and $\geq r_{.975}$
Normal	+.016	15	.0255	.0258	.0513
Rectangular	-.019	15	.0248	.0260	.0507
Leptokurtic	-.004	15	.0244	.0227	.0471
Slightly Skewed	-.010	15	.0220	.0256	.0476
Markedly Skewed	-.012	15	.0208	.0299	.0507
Normal	+.016	30	.0218	.0255	.0473
Rectangular	-.019	30	.0226	.0228	.0454
Leptokurtic	-.004	30	.0219	.0223	.0442
Slightly Skewed	-.010	30	.0224	.0274	.0498
Markedly Skewed	-.012	30	.0213	.0303	.0516
Normal	+.016	90	.0259	.0229	.0488
Rectangular	-.019	90	.0252	.0225	.0477
Leptokurtic	-.004	90	.0188	.0158	.0345
Slightly Skewed	-.010	90	.0199	.0267	.0465
Markedly Skewed	-.012	90	.0236	.0248	.0484

correlations in the separate tails did not deviate from the expected .0050 by more than .0010; the proportion in both tails combined did not deviate from the expected .0100 by more than .0016 even for samples as small as size fifteen.

Table 28

Proportions of Significant Coefficients in Sampling Distributions from Populations Having Approximately No Correlation ($\alpha = .01$)

Population	Population Correlation	N	$\leq r_{.005}$	$\geq r_{.995}$	$\leq r_{.005}$ and $\geq r_{.995}$
Normal	+.016	15	.0051	.0059	.0110
Rectangular	-.019	15	.0048	.0065	.0113
Leptokurtic	-.004	15	.0053	.0053	.0106
Slightly Skewed	-.010	15	.0052	.0063	.0115
Markedly Skewed	-.012	15	.0036	.0078	.0114
Normal	+.016	30	.0053	.0054	.0106
Rectangular	-.019	30	.0058	.0048	.0105
Leptokurtic	-.004	30	.0051	.0043	.0094
Slightly Skewed	-.010	30	.0046	.0068	.0114
Markedly Skewed	-.012	30	.0031	.0071	.0103
Normal	+.016	90	.0060	.0056	.0116
Rectangular	-.019	90	.0034	.0034	.0068
Leptokurtic	-.004	90	.0030	.0030	.0060
Slightly Skewed	-.010	90	.0041	.0071	.0113
Markedly Skewed	-.012	90	.0038	.0068	.0105

The obtained sampling distributions from the control population appeared very slightly leptokurtic in comparison with their theoretical distributions. In comparing a typical leptokurtic distribution with a normal distribution, the number of coefficients in the central portion of the curve would be greater with fewere coefficients on both sides of the mean and with a greater number of coefficients in the extreme tails. This agreed with the obtained discrepancies which were in the direction of too few significant coefficients for sample sizes 30 and 90 at the .05 level of significance and too many significant coefficients at the .01 level of significance. The discrepancies could be expected to become more pronounced with increased levels of significance.

The close correspondence of the obtained and theoretical sampling distributions where the population correlation was approximately zero and the score distribution approximately normal was interpreted as providing assurance of the adequacy of the sampling procedures used in this study. In particular, the effects of using discrete, finite populations in place of the theoretical continuous and infinite populations did not appear to materially influence the results. Nor did sampling procedures, as such, appear to cause deviation of the empirically determined distributions from the theoretical expectations.

Rectangular Population

The empirical sampling distributions based on rectangular populations conformed very closely to their theoretical distributions. However, the Kolmogorov-Smirnov tests indicated a significant departure between the two for sample size fifteen. Although the probability of such discrepancy occurring by chance alone was less than .05, the magnitude of the discrepancy was small (See Graphs 2, 7, and 12 of Appendix C). When relative proportion of obtained significant correlations was compared with the theoretical .0250 in each tail at the

.05 level of significance, the greatest discrepancy was .0025; the proportion in both tails combined did not deviate from the expected .0500 by more than .0046 for any one sample size. At the .01 level of significance the proportions of obtained significant correlations did not deviate from the expected .0050 by more than .0016 in either tail; the proportion in both tails combined did not deviate from the expected .0100 by more than .0032 for any one sample size.

All considered, there was little evidence that sampling from a rectangular population produced sampling distributions for the product moment correlation coefficient markedly different from those based on a normal population. Visual inspection of the appropriate graphs in Appendix C even indicates that such discrepancies as were obtained between the obtained and theoretical sampling distribution were very similar to those found between the theoretical and obtained distributions for the normal population.

Leptokurtic Population

The empirically determined sampling distributions based on samples of size 15 and 90 deviated significantly from their theoretical distributions at the .05 level of significance. For samples of size thirty no such significant deviation was noted even at the .20 level of significance. However, inspection of Graphs 3, 8, and 13 of Appendix C indicated that all three obtained sampling distributions deviated from the theoretical by having too many near-zero correlations and too few extremely large or extremely small values. When relative incidence of obtained significant correlations was compared with the theoretical .0250 in each tail at the .05 level of significance, the greatest discrepancy was .0092; for both tails combined the greatest discrepancy from the theoretical .0500 was .0155. At the .01 level of significance the largest discrepancy between the proportion of obtained significant correlations and the theoretical .0050 in each tail was .0020; for both tails combined the greatest discrepancy from the theoretical .0100 was .0040. At both levels of significance the size of the discrepancy increased as the sample size increased.

While the obtained sampling distributions based on a leptokurtic population deviated from their theoretical distributions in the same fashion as did those based on the normal population, the magnitude of the discrepancies was somewhat greater. For samples of size ninety significant correlations were obtained only .69 as often as expected at the .05 level of significance and .60 as often at the .01 level. It appeared that there was some evidence for the conclusion that the sampling distribution of correlation coefficients based on a leptokurtic population tended to be leptokurtic also and that the tendency increased with increased sample size.

Skewed Populations

The obtained sampling distributions based on markedly skewed and slightly skewed populations appeared to bear very similar relationships to their theoretical distributions. Inspection of Graphs 4, 5, 9, 10, 14, and 15 of Appendix C indicated very close correspondence between the obtained and theoretical distributions. Kolmogorov-Smirnov tests on the six comparisons indicated no significant departure from the theoretical distributions at the .20 level of significance except for samples of size fifteen based on markedly skewed population. The departure in that case was one which could occur between ten and fifteen percent of the time by chance alone. As indicated in tables 27 and 28 the proportions of significant correlations approximated the theoretical .0500 and .0100 very closely when the proportions in both tails of the distributions were combined. In no case did the proportion differ from expectation by more than .0035 at the .05 level of significance or by more than .0015 at the .01 level of significance. However, there was a consistent tendency for there to be too few significant negative correlations and too many significant positive correlations.

There was little evidence to suggest that the sampling distributions of correlation coefficients based on skewed data differed from their theoretical distributions. Only in comparison of relative incidence of extreme values did it appear that there may have been a tendency for the sampling distributions based on positively skewed populations to be positively skewed also.

Populations Having Substantial Correlations

The results of the Kolmogorov-Smirnov tests comparing the obtained and theoretical sampling distributions for the populations having substantial correlation are listed in Table 29. The proportions of coefficients in the obtained sampling distributions which were more extreme

Table 29

Results of Kolmogorov-Smirnov Tests Comparing Obtained Sampling Distributions from Populations Having Substantial Correlation with Their Theoretical Sampling Distributions

Population	Population Correlation	N	Number of Samples	Absolute Maximum Deviation
Normal	+.835	15	15,984	.0395*
Rectangular	+.831	15	15,984	.0933*
Leptokurtic	+.809	15	15,984	.0684*
Slightly Skewed	+.826	15	15,984	.0457*
Markedly Skewed	+.826	15	15.984	.0619*
Normal	+.835	30	7,992	.0286*
Rectangular	+.831	30	7,992	.0852*
Leptokurtic	+.809	30	7,992	.0760*
Slightly Skewed	+.826	30	7,992	.0363*
Markedly Skewed	+.826	30	7,992	.0518*
Normal	+.835	90	2,664	.0112
Rectangular	+.831	90	2,664	.0820*
Leptokurtic	+.809	90	2,664	.0856*
Slightly Skewed	+.826	90	2,664	.0420*
Markedly Skewed	+.826	90	2,664	.0451*

*Significant at $\alpha = .01$

than $r_{.005}$, $r_{.025}$, $r_{.975}$, and $r_{.995}$ are listed in tables 30 and 31.

Normal (Control) Population

The sampling distributions for sample sizes 15 and 30 from the control population having substantial correlation were significantly

Table 30

Proportions of Significant Coefficients in Sampling Distributions from Populations Having Substantial Correlation ($\alpha = .05$)

Population	Population Correlation	N	$\leq r_{.025}$	$\geq r_{.975}$	$\leq r_{.025}$ and $\geq r_{.975}$
Normal	+.835	15	.0194	.0300	.0494
Rectangular	+.831	15	.0054	.0134	.0188
Leptokurtic	+.809	15	.0389	.0681	.1070
Slightly Skewed	+.826	15	.0255	.0405	.0661
Markedly Skewed	+.826	15	.0297	.0515	.0811
Normal	+.835	30	.0211	.0288	.0499
Rectangular	+.831	30	.0044	.0085	.0129
Leptokurtic	+.809	30	.0533	.0763	.1296
Slightly Skewed	+.826	30	.0310	.0439	.0749
Markedly Skewed	+.826	30	.0333	.0488	.0821
Normal	+.835	90	.0221	.0252	.0473
Rectangular	+.831	90	.0038	.0064	.0101
Leptokurtic	+.809	90	.0732	.0830	.1562
Slightly Skewed	+.826	90	.0312	.0417	.0728
Markedly Skewed	+.826	90	.0375	.0473	.0848

different from their theoretical sampling distributions at the .01 level of significance as indicated by Kolmogorov-Smirnov tests. The sampling distribution for sample size ninety was not significantly different at the .20 level of significance. The proportions of significant coefficients in the sampling distributions approximated the expected

Table 31

Proportions of Significant Coefficients in Sampling Distributions from Populations Having Substantial Correlation ($\alpha = .01$)

Population	Population Correlation	N	$\leq r_{.005}$	$\geq r_{.995}$	$\leq r_{.005}$ and $\geq r_{.995}$
Normal	+.835	15	.0038	.0076	.0114
Rectangular	+.831	15	.0006	.0023	.0029
Leptokurtic	+.809	15	.0116	.0246	.0362
Slightly Skewed	+.826	15	.0065	.0112	.0177
Markedly Skewed	+.826	15	.0077	.0143	.0220
Normal	+.835	30	.0043	.0065	.0108
Rectangular	+.831	30	.0006	.0009	.0015
Leptokurtic	+.809	30	.0185	.0285	.0470
Slightly Skewed	+.826	30	.0123	.0116	.0239
Markedly Skewed	+.826	30	.0085	.0110	.0195
Normal	+.835	90	.0053	.0049	.0101
Rectangular	+.831	90	.0000	.0008	.0008
Leptokurtic	+.809	90	.0263	.0357	.0619
Slightly Skewed	+.826	90	.0131	.0116	.0248
Markedly Skewed	+.826	90	.0068	.0139	.0206

proportions when the negative and positive tails were summed. However, the proportions in the positive tails were consistently greater than in the negative tails except for sample size ninety at the .01 level of significance. The discrepancies decreased as sample size increased. The ratios of the proportions of significant coefficients in the positive and negative tailes were about 2 to 1 for sample size fifteen at the .01 level of significance, about 3 to 2 for sample size thirty at the .01 level of significance and sample sizes 15 to 30 at the .05 level of significance, and about 1 to 1 for sample size ninety at the .05 and .01 levels of significance. In regard to the tails the sampling distribution for sample size ninety tended to approximate the theoretical sampling distribution, whereas the sampling distributions for sample sizes 15 and 30 did not.

The deviation of the obtained sampling distributions from the theoretical ones based on the assumption of normality of distribution for Fishers' z statistic cast doubt on the validity of that assumption for values of rho as great as .835. For the purpose of determining the effect of nonnormality of population distributions on the sampling distributions of the product moment coefficient, the obtained versus theoretical comparisons seemed somewhat inappropriate. Consequently, the obtained sampling distributions based on nonnormal populations were compared also with the obtained sampling distributions based on the normal population.

Rectangular Population

The empirical sampling distributions for samples of size 15, 30, and 90 deviated from their theoretical distributions at the .01 level of significance according to the Kolmogorov-Smirnov tests. Inspection of Graphs 17, 22, and 27 of Appendix C showed that these departures were in the same direction as in the normal control population, but the effect was much more pronounced. The critical regions of the obtained distributions based on the rectangular population in no case contained more than .447 of the number of significant coefficients obtained from the normal population. The disproportionately low incidence of significant coefficients became more extreme for increased sample size and for lower level of significance. In one tail of the sampling distribution based on samples of size ninety there were no coefficients found as low as the theoretical value of $r_{.005}$.

All available evidence indicated that the sampling distributions of product moment coefficients based on a rectangular population were much more leptokurtic than either their theoretical distributions or the obtained distributions based on a normal population. At the .05 level of significance for both tails combined, the empirical distributions contained between .381 and .214 of the number of significant correlations found in the empirical distributions based on the normal population. At the .01 level of significance the corresponding proportions were .254 and .079. Rectangularity appeared to have a pronounced effect on the sampling distributions of the product moment correlation coefficient.

Leptokurtic Population

Obtained sampling distributions based on leptokurtic populations deviated from their theoretical distributions at the .01 level of significance using the Kolmogorov-Smirnov statistic. The proportions of significant coefficients in all critical regions were higher than anticipated from the theoretical distributions and from two to more than seven times as great as those obtained from the normal population. The disproportionately high incidence of significant coefficients became more pronounced for increased sample size and for lower level of significance. At the .05 level of significance for both tails combined, there were 2.166 to 3.302 times the number of significant coefficients obtained from the normal population. At the .01 level of significance the corresponding proportions were 3.175 and 6.129.

There was strong evidence that the sampling distributions of product moment correlations based on a leptokurtic population were significantly more platykurtic than either their theoretical distributions or the obtained sampling distributions based on a normal population (See Graphs 18, 23, and 28 of Appendix C). This was reflected in the incidence of extremely large and extremly small correlation coefficients in greater frequency than anticipated from the theoretical and control distributions.

Skewed Populations

Kolmogrov-Smirnov tests showed significant discrepancies between the obtained sampling distributions based on skewed populations and their theoretical distributions at the .01 level of significance. Inspection of Graphs 19, 20, 24, 25, 29, and 30 of Appendix C show that these discrepancies were of the same nature and magnitude for populations skewed either slightly or markedly. In all cases the critical regions contained a greater proportion of significant correlations than expected from the theoretical distributions and a still greater proportion than found in the empirical distributions based on a normal population. In all cases there were between 1.283 and 2.860 times the number of significant coefficients observed in the separate tails of the sampling distributions based on the control population. In nearly all cases the lower critical region contained proportionately more significant correlations than the upper critical region at the .05 level of significance. At the .01 level of significance the lower tail generally contained fewer significant coefficients but, compared to the proportions in distributions based on the normal population, their relative incidence was greater.

General Conclusions

Several general observations seemed warranted on the basis of the data from this study. They are:

1. When there was essentially no correlation in the population, the shape of the sampling distributions for the product moment correlation coefficients did not vary markedly as a function of nature or extent of nonnormality in the bivariate distribution. In general, these obtained sampling distributions conformed very closely to their theoretical distributions.
2. When there was substantial correlation in the population, the shape of the sampling distribution based on a bivariate normal population departed markedly from its theoretical distribution. Fishers' z transformation was apparently ineffective in providing a normally distributed statistic for values of rho as great as .835.
3. When there was substantial correlation in the population, the sampling distributions based on nonnormal populations deviated from their theoretical distributions and from

the distribution empirically determined for a bivariate normal population. The nature and direction of the deviation was specific to the type of nonnormality in the population; some sampling distributions appearing leptokurtic and others being platykurtic.

4. Discrepancies between proportions of theoretical and obtained significant correlations were relatively greater at the .01 level of significance than at the .05 level. Some of the obtained sampling distributions contained too many correlation coefficients in the extreme tails and too few coefficients in the less extreme sections of the tails. In other sampling distributions the reverse was true. Due to the inconsistency of the effects of nonnormality on the tails of the sampling distribution, it did not appear possible to adjust for nonnormality by modifying the level of significance in either direction.
5. When there was substantial correlation in the populations, increased sample size appeared to accentuate the discrepancy between the obtained sampling distributions and either their theoretical distributions or the empirical distribution based on a bivariate normal population.
6. Sampling distributions based on skewed populations yielded disproportionate numbers of significant correlations in the two tails of the distribution whether there was correlation in the populations or not. Where rho was

approximately zero, the two-tailed tests yielded approximately the expected proportion of extreme coefficients, but one-tailed tests based on the same distribution did not.

Relationship to Findings in Previous Research

In many respects the findings in the present study and the findings in previous research were similar. The previous studies generally found agreement for goodness of fit between the obtained sampling distributions and their theoretical sampling distributions when the population correlation was essentially zero. This was also true in the present study. Rider[2] found general agreement between the sampling distribution from a rectangular population with zero correlation and its expected distribution as indicated by a Chi-square test of goodness of fit. Dunlap[3] obtained a slightly leptokurtic sampling distribution from a rectangular population having zero correlation. In the present study the sampling distributions from the rectangular population having approximately no correlation were also slightly leptokurtic; however, the sampling distributions from the normal population having approximately no correlation were also slightly leptokurtic. Baker[4] obtained a skewed sampling distribution from a skewed population having zero correlation. The sampling distribution was skewed in the same direction in which the population was skewed. This corresponded to the findings in the current study.

---

[2] Paul R. Rider, "On the Distribution of the Correlation Coefficient in Small Samples," Biometrika, XXIV, Parts 3 and 4 (1932), 382-403.

[3] Hilda Frost Dunlap, "An Empirical Determination of the Distribution of Means, Standard Deviations and Correlation Coefficients Drawn from Rectangular Populations," The Annals of Mathematical Statistics, II, No. 1 (1931), 66-81.

[4] George A. Baker, "The Significance of the Product-Moment Coefficient of Correlation with Special Reference to the Character of the Marginal Distributions," Journal of the American Statistical Association, XXV, No. 172 (1930), 387-396.

In some instances the findings in the present study and the findings in previous research were dissimilar. Pearson[5] obtained approximately the expected number of significant coefficients in two sampling distributions from a skewed population having zero correlation. He concluded, therefore, that the effect of the skewness of the population had a negligible effect upon the sampling distributions. Pearson had pooled the coefficients from the negative and positive tails. In the present study the sampling distributions from the skewed populations having approximately no correlation contained approximately the expected proportions of significant coefficients when the negative and positive tails were pooled. However, the proportions in the negative and positive tails were not equal. Pearson[6] obtained too many significant coefficients in a sampling distribution for sample size ten and approximately the expected number in a sampling distribution for sample size twenty from a leptokurtic population having zero correlation. The result of too many significant coefficients for sample size ten was not in agreement with the findings of the present study. However, the fact that the sampling distribution for sample size ten in Pearson's study consisted of only 250 coefficients reduces the confidence which may be placed in the results.

---

[5] Egon S. Pearson, "The Test of Significance for the Correlation Coefficient," Journal of the American Statistical Association, XXVI, No. 174 (1931), 128-134.

[6] Ibid.

CHAPTER V

SUMMARY

There has not been unanimous agreement among researchers using product moment correlation as to the necessity of assuming the score distribution is bivariate normal. The assumption of normality need not be met if the product moment correlation coefficient as a descriptive statistic and as an index of linear relationship is required. However, if a statistical inference is to be made regarding a population product moment correlation coefficient, the assumptions which entered into the derivation of the sampling distributions theoretically must be met. The earlier works on product moment correlation dealt with bivariate normal populations. Sampling distributions were deduced for the product moment correlation coefficient assuming the samples were randomly drawn from normal populations. The purpose of the study was to investigate by empirical means the effects of nonnormality upon some sampling distributions of the product moment correlation coefficient.

Procedure

Ten populations of size ten thousand were established. A population having approximately no correlation and a population having substantial correlation were established for each of the following forms: normal, rectangular, leptokurtic, slightly skewed, and markedly skewed. The ten populations were punched onto one set of ten thousand IBM cards. The ten thousand cards were placed in random order by ordering four columns of random digits. The samples were obtained by repeatedly counting off the desired number of cards. The deck was sampled without replacement and to exhaustion. The deck was reordered by using another set of four columns of random digits and samples counted off. This was repeated twenty-four times in order to obtain the desired number of samples. This necessitated having five decks of ten thousand cards each with the sixty columns of data being identical for each deck leaving a total of one hundred columns for random digits. Sample sizes 15, 30, and 90 were obtained concurrently with the first two samples of size fifteen comprising the first sample of size thirty and the first three samples of size thirty comprising the first sample of size ninety. Correlation coefficients were calculated for the samples and sampling distributions for samples sizes 15, 30, and 90 were established for each of the ten populations. The sampling distributions consisted of 15,984 coefficients for sample size fifteen, 7,992 coefficients for sample size thirty, and 2,664 coefficients for sample size ninety. The sampling and calculating were done by means of the IBM 650 Data Processing System.

The obtained sampling distributions were compared with their theoretical sampling distributions and with the obtained sampling distributions from the normal populations. Fisher's transformations were used in establishing the theoretical sampling distributions. The obtained and theoretical sampling distributions were tested for overall goodness of fit by means of Kolmogorov-Smirnov tests. Critical values were established for $r_{.005}$, $r_{.025}$, $r_{.975}$, and $r_{.995}$ and frequency tabulations made of the number of significant coefficients in each of the obtained sampling distributions.

Results and Conclusions

The obtained sampling distributions from the control population having approximately no correlation tended to approximate their theoretical sampling distributions. Therefore, the obtained sampling distributions from the experimental populations having approximately no correlation were compared with their theoretical sampling distributions and to the obtained sampling distributions from the control population. The obtained sampling distribution for sample size ninety from the control population having substantial correlation tended to approximate its theoretical sampling distribution; however, the obtained sampling distributions for sample sizes 15 and 30 from the control population having substantial correlation did not tend to approximate their theoretical sampling distributions. Therefore, the obtained sampling distributions from the experimental populations having substantial correlation were compared primarily with the obtained sampling distributions from the control population.

When the population correlation approximated zero, the obtained sampling distributions tended to conform to their theoretical expectations. The distributions from the rectangular population conformed very closely to their expected distributions. The distributions from the leptokurtic population tended to be slightly leptokurtic; however, they also conformed quite closely to the expected distributions. There was a slight tendency for the distributions from the skewed populations to be skewed and in the same direction. Overall, they too tended to conform to the expected distributions.

When the population correlation was substantial, the obtained sampling distributions deviated from their theoretical distributions and from the obtained distributions from the control population. The distributions from the rectangular population were extremely leptokurtic. The number of significant coefficients was drastically reduced. The distributions from the leptokurtic population were more platykurtic than either their theoretical distributions or the obtained distributions from the control population. The number of significant coefficients was considerably greater than in the distributions from the control population. The distributions from the skewed populations were skewed in comparison with their theoretical distributions. They contained considerably more significant coefficients than either their theoretical distributions or the obtained distributions from the control population.

Several general observations seemed warranted on the basis of the data from this study. They are:

1. When there was essentially no correlation in the population, the shape of the sampling distributions for the product moment correlation coefficients did not vary markedly as a function of nature or extent of nonnormality in the bivariate distribution. In general, these obtained sampling distributions conformed very closely to their theoretical distributions.
2. When there was substantial correlation in the population, the shape of the sampling distribution based on a bivariate normal population departed markedly from its theoretical distribution. Fishers' z transformation was apparently

ineffective in providing a normally distributed statistic for values of rho as great as .835.

3. When there was substantial correlation in the population, the sampling distributions based on nonnormal populations deviated from their theoretical distributions and from the distribution empirically determined for a bivariate normal population. The nature and direction of the deviation was specific to the type of nonnormality in the population; some sampling distributions appearing leptokurtic and others being platykurtic.
4. Discrepancies between proportions of theoretical and obtained significant correlations were relatively greater at the .01 level of significance than at the .05 level. Some of the obtained sampling distributions contained too many correlation coefficients in the extreme tails and too few coefficients in the less extreme sections of the tails. In other sampling distributions the reverse was true. Due to the inconsistency of the effects of nonnormality on the tails of the sampling distribution, it did not appear possible to adjust for nonnormality by modifying the level of significance in either direction.
5. When there was substantial correlation in the populations, increased sample size appeared to accentuate the discrepancy between the obtained sampling distributions and either their theoretical distributions or the empirical distribution

based on a bivariate normal population.

6. Sampling distributions based on skewed populations yielded disproportionate numbers of significant correlations in the two tails of the distribution whether there was correlation in the populations or not. Where rho was approximately zero, the two-tailed tests yielded approximately the expected proportion of extreme coefficients, but one-tailed tests based on the same distribution did not.

The results of the investigation of the effects of nonnormality upon the sampling distributions of the product moment correlation coefficient indicated that nonnormality had an effect and that it was of such a magnitude that it should be taken into account in dealing with statistical tests of inference involving product moment correlation. More research is needed, especially in regard to other types of nonnormality and also in regard to populations not having identical marginal frequencies.

BIBLIOGRAPHY

Baker, George A. "The Significance of the Product-Moment Coefficient of Correlation with Special Reference to the Character of the Marginal Distributions," Journal of the *American Statistical Association*, XXV (1930), 387-396.

Binder, Arnold. "Consideration of the Place of Assumptions in Correlational Analysis," *The American Psychologist* XIV (1959), 504-510.

Chesire, Leone, Oldis, Elena, and Pearson, Egon S. "Further Experiments on the Sampling Distribution of the Correlation Coefficient," *Journal of the American Statistical Association*, XXVII (1932), 121-128.

David, F. N. *Tables of the Ordinates and Probability Integral of the Distribution of the Correlation Coefficient in Small Samples.* Cambridge, England: The University Press, 1938.

Dunlap, Hilda Frost. "An Empirical Determination of Means, Standard Deviations and Correlation Coefficients Drawn from Rectangular Populations," *The Annals of Mathematical Statistics*, II (1931), 66-81.

Fisher, R. A. "On the 'Proba[illegible] a Coefficient of Correlation Deduced from a Small Sample," *Metron*, I (1921), 1-32.

Kendall, Maurice G. *The Advanced Theory of Statistics*, I. 3rd ed. London: Charles Griffin & Company Limited, 1947.

Li, Jerome C. R. *Introduction to Statistical Inference.* Ann Arbor: Edwards Brothers, Inc., 1957.

McNemar, Quinn. *Psychological Statistics.* 2nd ed. New York: John Wiley & Sons, Inc., 1955.

Mulholland, John H. "Comment on 'The Needless Assumption of Normality in Pearson's r.'" *The American Psychologist*, XIII (195), 544-545.

Nefzger, M. D., and Drasgow, James. "The Needless Assumption of Normality in Pearson's r," *The American Psychologist*, XII (1957), 623-625.

Norton, Dee Wixwell. "An Empirical Investigation of the Effects of Nonnormality and Heterogeneity upon the F-test of Analysis of Variance." Unpublished Ph.D. dissertation, State University of Iowa, 1952.

Pearson, Egon S. "Some Notes on Sampling Tests with Two Variables," *Biometrika*, XXI (1929), 337-360.

BIBLIOGRAPHY

BIBLIOGRAPHY

Baker, George A. "The Significance of the Product-Moment Coefficient of Correlation with Special Reference to the Character of the Marginal Distributions," Journal of the *American Statistical Association*, XXV (1930), 387-396.

Binder. Arnold, "Consideration of the Place of Assumptions in Correlational Analysis," *The American Psychologist* XIV (1959), 504-510.

Chesire, Leone, Oldis, Elena, and Pearson, Egon S. "Further Experiments on the Sampling Distribution of the Correlation Coefficient," *Journal of the American Statistical Association*, XXVII (1932), 121-128.

David, F. N. *Tables of the Ordinates and Probability Integral of the Distribution of the Correlation Coefficient in Small Samples.* Cambridge, England: The University Press, 1938.

Dunlap, Hilda Frost. "An Empirical Determination of Means, Standard Deviations and Correlation Coefficients Drawn from Rectangular Populations," *The Annals of Mathematical Statistics*, II (1931), 66-81.

Fisher, R. A. "On the 'Probable Error' of a Coefficient of Correlation Deduced from a Small Sample," *Metron*, I (1921), 3-32.

Kendall, Maurice G. *The Advanced Theory of Statistics*, I. 3rd ed. London: Charles Griffin & Company Limited, 1947.

Li, Jerome C. R. *Introduction to Statistical Inference.* Ann Arbor: Edwards Brothers, Inc., 1957.

McNemar, Quinn. *Psychological Statistics.* 2nd ed. New York: John Wiley & Sons, Inc., 1955.

Milholland, John E. "Comment on 'The Needless Assumption of Normality in Pearson's r,'" *The American Psychologist*, XIII (1958), 544-546.

Nefzger, M. D., and Drasgow, James. "The Needless Assumption of Normality in Pearson's r," *The American Psychologist*, XII (1957), 623-625.

Norton, Dee Wiswall, "An Empirical Investigation of the Effects of Nonnormality and Heterogeneity upon the F-test of Analysis of Variance." Unpublished Ph.D. dissertation, State University of Iowa, 1952.

Pearson, Egon S. "Some Notes on Sampling Tests with Two Variables," *Biometrika*, XXI (1929), 337-360.

________. "The Test of Significance for the Correlation Coefficient," *Journal of the American Statistical Association*, XXVI (1931), 128-134.

________. "The Test of Significance for the Correlation Coefficient: Some Further Results," *Journal of the American Statistical Association*, XXVII (1932), 424-426.

Pearson, Karl. "Notes on the History of Correlation," *Biometrika*, XIII (1920), 25-45.

Peters, Charles C., and Van Voorhis, Walter R. *Statistical Procedures and Their Mathematical Bases*. 1st ed. New York: McGraw-Hill Book Company, Inc., 1940.

Rand Corporation. *A Million Random Digits with 100,000 Normal Deviates*. Glencoe: Free Press, 1955.

Rider, Paul R. "On the Distribution of the Correlation Coefficient in Small Samples," *Biometrika*, XXIV (1932), 382-403.

Siegel, Sidney. *Nonparametric Statistics for the Behavioral Sciences*. New York: McGraw-Hill Book Company, Inc., 1956.

Walker, Helen M., and Lev, Joseph. *Statistical Inference*. New York: Henry Holt and Company, 1953.

APPENDIX

APPENDIX A

TABLES OF BIVARIATE POPULATIONS

Table 32

Bivariate Normal Population Having a Correlation of .000

X_6 Value	X_1 Value																				Marginal Frequency
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
19										1	1										2
18								1	1	1	1	1	1								6
17						1	1	2	3	4	3	3	2	1	1						21
16				1	1	2	4	6	8	10	10	8	6	4	2	1	1				64
15			1	1	3	5	10	15	21	24	24	21	16	10	5	3	1				160
14			1	2	5	12	21	33	44	51	51	44	33	21	12	5	2	1			338
13			1	4	10	21	38	60	81	93	93	81	60	38	21	10	4	1	1		617
12		1	2	6	16	33	60	93	126	147	147	126	93	60	33	16	6	2	1		968
11		1	3	8	21	44	81	126	171	198	198	171	126	81	44	21	8	3	1		1,306
10	1	1	4	10	24	51	93	147	198	230	230	198	147	93	51	24	10	4	1	1	1,518
9	1	1	4	10	24	51	93	147	198	230	230	198	147	93	51	24	10	4	1	1	1,518
8		1	3	8	21	44	81	126	171	198	198	171	126	81	44	21	8	3	1		1,306
7		1	2	6	16	33	60	93	126	147	147	126	93	60	33	16	6	2	1		968
6		1	1	4	10	21	38	60	81	93	93	81	60	38	21	10	4	1			617
5			1	2	5	12	21	33	44	51	51	44	33	21	12	5	2	1			338
4				1	3	5	10	16	21	24	24	21	15	10	5	3	1	1			160
3				1	1	2	4	6	8	10	10	8	6	4	2	1	1				64
2						1	1	2	3	3	4	3	2	1	1						21
1								1	1	1	1	1	1								6
0										1	1										2
Marginal Frequency	2	7	23	64	160	338	616	967	1,306	1,517	1,517	1,306	967	616	338	160	64	23	7	2	10,000

Table 33

Bivariate Rectangular Population Having a Correlation of .000

X_7 Value	X_2 Value																				Marginal Frequency
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
19	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	500
18	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	500
17	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	500
16	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	500
15	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	500
14	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	500
13	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	500
12	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	500
11	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	500
10	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	500
9	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	500
8	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	500
7	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	500
6	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	500
5	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	500
4	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	500
3	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	500
2	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	500
1	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	500
0	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	500
Marginal Frequency	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	10,000

Table 34

Bivariate Leptokurtic Population Having a Correlation of .000

X_8 Value	X_3 Value																				Marginal Frequency
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
19										1	1										2
18										1	1										2
17										1	1										2
16									1	1	1										3
15									2	4	4	2									12
14								1	4	9	9	4	1								28
13							1	4	13	32	32	13	4	1							100
12						1	4	13	48	115	115	48	13	4	1						362
11			1		2	4	13	48	172	416	417	172	48	13	4	2	1				1,313
10	1	1	1	1	4	10	32	115	417	1,006	1,006	417	115	32	10	4	1	1	1	1	3,176
9	1	1	1	1	4	10	32	115	417	1,006	1,006	417	115	32	10	4	1	1	1	1	3,176
8				1	2	4	13	48	172	417	416	172	48	13	4	2		1			1,313
7						1	4	13	48	115	115	48	13	4	1						362
6							1	4	13	32	32	13	4	1							100
5								1	4	9	9	4	1								28
4									2	4	4	2									12
3										1	1	1									3
2										1	1										2
1										1	1										2
0										1	1										2
Marginal Frequency	2	2	3	3	12	30	100	362	1,313	3,173	3,173	1,313	362	100	30	12	3	3	2	2	10,000

Table 35

Bivariate Slightly Skewed Population Having a Correlation of -.005

X_9 Value	X_4 Value 0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	Marginal Frequency
17			1		1	1	1														4
16			1	1	1	1	1	1	1	1											8
15			1	2	2	3	2	2	1	1	1										15
14			2	4	5	6	5	4	3	2	1	1									33
13		1	4	7	10	11	10	8	6	4	2	1	1								65
12		2	7	13	18	20	18	15	11	7	4	3	2	1							121
11		3	11	23	33	36	33	26	19	12	8	4	3	1	1						213
10		5	19	39	56	61	56	44	32	21	13	8	4	2	1	1					362
9	1	8	31	63	89	98	89	72	52	34	21	12	7	4	2	1	1				585
8	1	12	47	96	135	147	135	108	78	52	32	19	11	6	3	1	1	1			884
7	1	16	64	132	187	204	187	150	108	72	45	26	15	8	4	2	1	1	1		1,223
6	1	21	81	166	234	256	234	187	135	89	56	33	18	10	5	3	1	1	1	1	1,533
5	2	22	88	181	256	280	256	205	147	98	61	36	20	11	6	3	2	1			1,677
4	2	21	80	166	233	256	234	187	135	89	56	33	18	10	5	2	1				1,529
3	1	15	57	117	165	181	166	132	95	63	40	23	13	7	4	2	1				1,082
2	1	7	22	57	80	88	80	64	46	31	19	11	6	3	2	1	1				519
1		2	7	15	21	22	21	17	12	8	5	3	2	1	1						137
0			1	1	2	2	2	1	1	1											11
Marginal Frequency	10	135	524	1,063	1,528	1,673	1,530	1,223	882	585	364	213	120	64	34	16	9	4	2	1	10,000

Table 36

Bivariate Markedly Skewed Population Having a Correlation of -.004

X_{10} Value	X_5 Value 0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	Marginal Frequency
17	1	1	1	1	1																5
16	1	2	2	1	1	1	1	1													10
15	2	3	3	3	2	2	1	1	1	1	1										20
14	4	5	5	5	4	3	3	2	2	2	1	1	1								36
13	6	9	9	8	7	6	5	3	3	2	1	1	1								61
12	10	14	14	13	11	9	7	6	4	3	2	1	1	1							96
11	15	22	22	20	17	14	11	8	6	5	3	2	2	1	1						149
10	22	32	32	29	25	20	16	12	9	7	5	3	2	1	1	1					217
9	31	45	45	41	35	29	23	18	13	10	7	5	3	2	1	1					309
8	43	63	63	56	48	40	31	24	18	13	9	6	4	3	2	1	1				425
7	58	84	84	76	64	53	42	32	24	18	13	9	6	4	2	1	1				571
6	75	109	109	98	84	69	54	42	31	23	16	11	7	5	3	2	1				739
5	95	137	137	124	106	87	69	53	40	29	20	14	9	6	3	2	1				932
4	115	167	167	151	129	105	84	64	48	35	25	17	11	7	4	2	1	1			1,133
3	135	196	196	177	151	124	98	76	57	41	[illegible]	20	13	9	5	3	1	1	1		1,332
2	150	217	217	196	167	137	109	84	63	45	32	22	15	9	6	3	2	1	1	1	1,477
1	150	217	217	196	167	137	109	84	63	45	32	22	14	9	5	3	2	1	1		1,474
0	103	149	150	135	115	94	75	58	43	31	22	15	10	9	4	2	1	1			1,014
Marginal Frequency	1,016	1,472	1,473	1,330	1,134	930	738	568	425	309	218	149	98	62	37	21	11	5	3	1	10,000

Table 37

Bivariate Normal Population Having a Correlation of +.832

X_{11} Value	X_1 Value																				Marginal Frequency
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
21																1	1	1	1		4
20															1	2	3	3	2	1	12
19													1	2	5	8	8	5	1	1	31
18												1	3	10	18	19	13	5	2		71
17											1	4	15	32	40	31	14	5	1		143
16										1	5	20	49	72	66	38	13	3			267
15									1	5	24	67	114	121	78	31	8	1			450
14								1	4	24	78	154	189	142	66	19	3				680
13								3	20	78	178	256	225	121	40	8	1				930
12							2	15	67	178	297	302	189	72	18	2					1,142
11						1	10	49	154	297	351	256	114	32	5	1					1,270
10					1	5	32	114	256	351	297	154	49	10	1						1,270
9					2	8	72	189	302	297	178	67	15	2							1,142
8				1	8	40	121	225	256	178	78	20	3								930
7				3	19	66	142	189	154	78	24	4	1								680
6			1	8	31	78	121	114	67	24	5	1									450
5			3	13	38	66	72	49	20	5	1										267
4		1	5	14	31	40	32	15	4	1											143
3		2	5	13	19	18	10	3	1												71
2	1	1	5	8	8	5	2	1													31
1	1	2	3	3	2	1															12
0		1	1	1	1																4
Marginal Frequency	2	7	23	64	160	338	616	967	1,306	1,517	1,517	1,306	967	616	338	160	64	23	7	2	10,000

Table 38

Bivariate Rectangular Population Having a Correlation of +.834

X_{12} Value	X_2 Value																				Marginal Frequency
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
33																				3	3
32																			3	38	41
31																		3	38	38	79
30																	3	38	38	38	117
29																3	38	38	38	38	155
28															3	38	38	38	38	38	193
27														3	38	38	38	38	38	38	231
26													3	38	38	38	38	38	38	38	269
25												3	38	38	38	38	38	38	38	38	307
24											3	38	38	38	38	38	38	38	38	38	345
23										3	38	38	38	38	38	38	38	38	38	38	383
22									3	38	38	38	38	38	38	38	38	38	38	38	421
21								3	38	38	38	38	38	38	38	38	38	38	38	38	459
20							3	38	38	38	38	38	38	38	38	38	38	38	38	38	497
19						3	38	38	38	38	38	38	38	38	38	38	38	38	38	3	500
18					3	38	38	38	38	38	38	38	38	38	38	38	38	38	3		500
17				3	38	38	38	38	38	38	38	38	38	38	38	38	38	3			500
16			3	38	38	38	38	38	38	38	38	38	38	38	38	38	3				500
15		3	38	38	38	38	38	38	38	38	38	38	38	38	38	3					500
14	3	38	38	38	38	38	38	38	38	38	38	38	38	38	3						500
13	38	38	38	38	38	38	38	38	38	38	38	38	38	3							497
12	38	38	38	38	38	38	38	38	38	38	38	38	3								459
11	38	38	38	38	38	38	38	38	38	38	38	3									421
10	38	38	38	38	38	38	38	38	38	38	3										383
9	38	38	38	38	38	38	38	38	38	3											345
8	38	38	38	38	38	38	38	38	3												307
7	38	38	38	38	38	38	38	3													269
6	38	38	38	38	38	38	3														231
5	38	38	38	38	38	3															193
4	38	38	38	38	3																155
3	38	38	38	3																	117
2	38	38	3																		79
1	38	3																			41
0	3																				3
Marginal Frequency	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	10,000

Table 30

Bivariate Leptokurtic Population Having a Correlation of .813

X_{15} Value	X_9 Value 0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	Marginal Frequency
19																				1	1
18																		1	1	1	3
17											1						1	1	1		4
16										1	1				1	3	1	1			6
15										1	1	2	2	4	6	6	1				23
14										1	4	7	14	22	16	3					67
13										4	15	45	70	48	6						201
12									2	15	110	290	170	22	1						618
11									6	110	697	626	70	4							1,384
10								2	45	697	1,520	290	14								2,646
9								14	290	1,520	697	45	2								2,646
8							4	70	626	697	110	6									1,384
7						1	22	170	290	110	15	2									618
6						6	48	70	45	15	4										201
5					3	16	22	14	7	4	1										67
4				1	6	6	4	2	2	1	1										23
3			1	1	3	1				1	1										6
2		1	1	1						1	1										4
1	1	1	1							1											3
0	1																				1
Marginal Frequency	2	2	3	3	12	30	100	342	1,310	3,178	3,178	1,318	342	100	30	12	3	3	2	2	10,000

Table [illegible]

Bivariate Slightly Skewed Population Having a Correlation of [illegible]

X_{14} Value	X_1 Value																				Marginal Frequency
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
21																	[illegible]		1	1	[illegible]
20														[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1		[illegible]
19											[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]			[illegible]
18										[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]			[illegible]
17									[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]				[illegible]
16							[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]				[illegible]
15						[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]					[illegible]
14					[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]						[illegible]
13					[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]						[illegible]
12			[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]							[illegible]
11			[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]								[illegible]
10			[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]									[illegible]
9			[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]										[illegible]
8		[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]											[illegible]
7		[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]											[illegible]
6		[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]												[illegible]
5		[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]													[illegible]
4	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]														[illegible]
3	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]															[illegible]
2	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]																[illegible]
1	[illegible]	[illegible]	[illegible]	[illegible]																	[illegible]
0	[illegible]	[illegible]	[illegible]																		[illegible]
Marginal Frequency	10	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	64	34	18	9	4	2	1	10,000

Table 41

Bivariate Markedly Skewed Population Having a Correlation of +.632

	X_1 Value																				
X_{13} Value	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	Marginal Frequency
22																1					1
21													1	1	1	1	1				5
20											1	1	1	1	2	2	2	1	1	1	13
19										1	1	2	3	3	3	2	1	1	1		18
18								1	1	2	3	3	4	5	4	3	2	2	1		31
17							1	1	3	4	5	7	7	7	5	4	3	1			48
16						1	1	4	6	9	10	10	10	10	7	5	2				74
15					1	2	6	7	10	14	15	16	18	11	8	3					108
14				1	3	6	9	14	16	22	25	22	15	14	7						157
13			3	4	7	12	19	24	30	32	38	20	22	10							223
12		2	4	9	15	24	21	40	43	45	42	33	17								308
11	1	4	10	17	29	30	32	50	64	50	40	26									407
10	3	10	16	24	48	67	73	68	81	67	36										526
9	7	16	30	36	80	85	111	120	94	34											661
8	15	30	61	94	118	140	140	121	73												803
7	28	61	108	134	170	176	144	100													938
6	48	108	151	199	218	216	129														1,048
5	72	151	221	213	200	162															1,109
4	104	221	279	294	139																1,087
3	153	279	326	233																	990
2	193	328	256																		777
1	228	237																			468
0	178																				178
Marginal Frequency	1,026	1,478	1,473	1,330	1,134	950	796	568	433	309	216	140	98	62	37	21	11	5	3	1	10,000

APPENDIX B

TABLES OF SAMPLING DISTRIBUTIONS OF THE PRODUCT MOMENT CORRELATION COEFFICIENT

Table 42

Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Normal Population ($\rho = +.016$, N = 15)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
+.80 to +.85	1	1.6	.0001	.0001
+.75 to +.80	5	6.4	.0003	.0004
+.70 to +.75	18	17.6	.0011	.0011
+.65 to +.70	51	43.2	.0032	.0027
+.60 to +.65	101-1/2	83.1	.0064	.0052
+.55 to +.60	131	142.3	.0082	.0089
+.50 to +.55	191	223.8	.0119	.0140
+.45 to +.50	287	316.5	.0180	.0198
+.40 to +.45	395	428.4	.0247	.0268
+.35 to +.40	531-1/2	545.1	.0333	.0341
+.30 to +.35	667-1/2	663.3	.0418	.0415
+.25 to +.30	780	781.6	.0488	.0489
+.20 to +.25	879-1/2	885.5	.0550	.0554
+.15 to +.20	986-1/2	975.0	.0617	.0610
+.10 to +.15	1,079	1,042.2	.0675	.0652
+.05 to +.10	1,107-1/2	1,086.9	.0693	.0680
.00 to +.05	1,136-1/2	1,102.9	.0711	.0690
-.05 to .00	1,133	1,091.7	.0709	.0683
-.10 to -.05	1,136-1/2	1,056.5	.0711	.0661
-.15 to -.10	1,001	994.2	.0626	.0622
-.20 to -.15	921	911.1	.0576	.0570
-.25 to -.20	785	810.4	.0491	.0507
-.30 to -.25	689-1/2	701.7	.0431	.0439
-.35 to -.30	533	581.8	.0333	.0364
-.40 to -.35	439	469.9	.0275	.0294
-.45 to -.40	346-1/2	358.0	.0217	.0224
-.50 to -.45	247-1/2	262.1	.0155	.0164
-.55 to -.50	181-1/2	177.4	.0114	.0111
-.60 to -.55	103-1/2	111.9	.0065	.0070
-.65 to -.60	57	62.3	.0036	.0039
-.70 to -.65	40	32.0	.0025	.0020
-.75 to -.70	17-1/2	12.8	.0011	.0008
-.80 to -.75	4	3.2	.0003	.0002

Table 42 (continued)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
-.85 to -.80	0	1.6	.0000	.0001
Total	15,984	15,984.0	1.0002	1.0000

Table 43

Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Rectangular Population ($\rho = -.019$, N = 15)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
+.80 to +.85	3	1.6	.0002	.0001
+.75 to +.80	5	3.2	.0003	.0002
+.70 to +.75	15	12.8	.0009	.0008
+.65 to +.70	41	30.4	.0026	.0019
+.60 to +.65	66	60.7	.0041	.0038
+.55 to +.60	107-1/2	108.7	.0067	.0068
+.50 to +.55	162-1/2	174.2	.0102	.0109
+.45 to +.50	247-1/2	255.7	.0155	.0160
+.40 to +.45	311-1/2	353.2	.0195	.0221
+.35 to +.40	434-1/2	460.3	.0272	.0288
+.30 to +.35	504-1/2	578.6	.0316	.0362
+.25 to +.30	697-1/2	692.1	.0436	.0433
+.20 to +.25	821-1/2	804.0	.0514	.0503
+.15 to +.20	892-1/2	906.3	.0558	.0567
+.10 to +.15	965-1/2	987.8	.0604	.0618
+.05 to +.10	1,034-1/2	1,053.3	.0647	.0659
.00 to +.05	1,088-1/2	1,090.1	.0681	.0682
-.05 to .00	1,138	1,102.9	.0712	.0690
-.10 to -.05	1,187-1/2	1,088.5	.0743	.0681
-.15 to -.10	1,103-1/2	1,047.0	.0690	.0655
-.20 to -.15	994	979.8	.0622	.0613
-.25 to -.20	908	893.5	.0568	.0559
-.30 to -.25	813	786.4	.0509	.0492
-.35 to -.30	688	672.9	.0430	.0421
-.40 to -.35	517-1/2	553.0	.0324	.0346
-.45 to -.40	436-1/2	433.2	.0273	.0271
-.50 to -.45	296-1/2	324.5	.0185	.0203
-.55 to -.50	200-1/2	227.0	.0125	.0142
-.60 to -.55	151-1/2	145.5	.0095	.0091
-.65 to -.60	86	86.3	.0054	.0054
-.70 to -.65	37	44.8	.0023	.0028
-.75 to -.70	17-1/2	17.6	.0011	.0011

Table 43 (continued)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
-.80 to -.75	6	6.4	.0004	.0004
-.85 to -.80	4	1.6	.0003	.0001
-.90 to -.85	0	.0	.0000	.0000
-.95 to -.90	1	.0	.0001	.0000
Total	15,984	15,983.9	1.0000	1.0000

Table 44

Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Leptokurtic Population ($\rho = -.004$, N = 15)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
+.80 to +.85	3	1.6	.0002	.0001
+.75 to +.80	1-1/2	4.8	.0001	.0003
+.70 to +.75	23-1/2	14.4	.0015	.0009
+.65 to +.70	40	35.2	.0025	.0022
+.60 to +.65	67	67.1	.0042	.0044
+.55 to +.60	94	121.5	.0059	.0076
+.50 to +.55	173-1/2	193.4	.0109	.0121
+.45 to +.50	253	281.3	.0158	.0176
+.40 to +.45	354-1/2	385.2	.0222	.0241
+.35 to +.40	447	495.5	.0280	.0310
+.30 to +.35	577	613.8	.0361	.0384
+.25 to +.30	748	730.5	.0468	.0457
+.20 to +.25	848-1/2	840.8	.0531	.0526
+.15 to +.20	944-1/2	935.1	.0591	.0585
+.10 to +.15	1,034	1,013.4	.0647	.0634
+.05 to +.10	1,189-1/2	1,069.3	.0744	.0669
.00 to +.05	1,183-1/2	1,098.1	.0740	.0687
-.05 to .00	1,149	1,099.7	.0719	.0688
-.10 to -.05	1,079-1/2	1,077.3	.0675	.0674
-.15 to -.10	1,060-1/2	1,024.6	.0663	.0641
-.20 to -.15	949	952.6	.0594	.0596
-.25 to -.20	829	858.3	.0519	.0537
-.30 to -.25	730	749.6	.0457	.0469
-.35 to -.30	625-1/2	634.6	.0391	.0397
-.40 to -.35	508	516.3	.0318	.0323
-.45 to -.40	364-1/2	399.6	.0228	.0250
-.50 to -.45	261-1/2	297.3	.0164	.0186
-.55 to -.50	199-1/2	203.0	.0125	.0127
-.60 to -.55	103-1/2	131.1	.0065	.0082
-.65 to -.60	78-1/2	75.1	.0049	.0047
-.70 to -.65	36	38.4	.0023	.0024
-.75 to -.70	21	16.0	.0013	.0010

Table 44 (continued)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
-.80 to -.75	7	4.8	.0004	.0003
-.85 to -.80	0	1.6	.0000	.0001
Total	15,984	15,980.9	1.0002	1.0000

Table 45

Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Slightly Skewed Population ($\rho = -.010$, N = 15)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
+.85 to +.90	2	.0	.0001	.0000
+.80 to +.85	4	1.6	.0003	.0001
+.75 to +.80	4	3.2	.0003	.0002
+.70 to +.75	16	14.4	.0010	.0009
+.65 to +.70	35	33.6	.0022	.0021
+.60 to +.65	75-1/2	65.5	.0047	.0041
+.55 to +.60	121-1/2	116.7	.0076	.0073
+.50 to +.55	168	187.0	.0105	.0117
+.45 to +.50	255	270.1	.0160	.0169
+.40 to +.45	365	370.8	.0228	.0232
+.35 to +.40	455	484.3	.0285	.0303
+.30 to +.35	571-1/2	596.2	.0358	.0373
+.25 to +.30	720-1/2	716.1	.0451	.0448
+.20 to +.25	836	826.4	.0523	.0517
+.15 to +.20	872-1/2	923.9	.0546	.0578
+.10 to +.15	1,026	1,003.8	.0642	.0628
+.05 to +.10	1,029	1,062.9	.0644	.0665
.00 to +.05	1,109	1,094.9	.0694	.0685
-.05 to .00	1,133	1,101.3	.0709	.0689
-.10 to -.05	1,143	1,082.1	.0715	.0677
-.15 to -.10	1,100-1/2	1,034.2	.0689	.0647
-.20 to -.15	1,036-1/2	963.8	.0648	.0603
-.25 to -.20	853-1/2	872.7	.0534	.0546
-.30 to -.25	755	762.4	.0472	.0477
-.35 to -.30	628	650.5	.0393	.0407
-.40 to -.35	523	532.3	.0327	.0333
-.45 to -.40	413	412.4	.0258	.0258
-.50 to -.45	293-1/2	306.9	.0184	.0192
-.55 to -.50	204	214.2	.0128	.0134
-.60 to -.55	114	135.9	.0071	.0085
-.65 to -.60	54	79.9	.0034	.0050
-.70 to -.65	39-1/2	40.0	.0025	.0025

Table 45 (continued)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
-.75 to -.70	16-1/2	17.6	.0010	.0011
-.80 to -.75	6-1/2	4.8	.0004	.0003
-.85 to -.80	5	1.6	.0003	.0001
Total	15,984	15,984.0	1.0002	1.0000

Table 46

Obtained and Theoretical Sampling Distributions of the Product moment Correlation Coefficient Based on Markedly Skewed Population ($\rho = -.012$, N = 15)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
+.80 to +.85	1	1.6	.0001	.0001
+.75 to +.80	14	3.2	.0009	.0002
+.70 to +.75	23	14.4	.0014	.0009
+.65 to +.70	38	32.0	.0024	.0020
+.60 to +.65	91-1/2	65.5	.0057	.0041
+.55 to +.60	135	115.1	.0084	.0072
+.50 to +.55	183-1/2	182.2	.0115	.0114
+.45 to +.50	292-1/2	266.9	.0183	.0167
+.40 to +.45	371-1/2	367.6	.0232	.0230
+.35 to +.40	454-1/2	479.5	.0284	.0300
+.30 to +.35	565	591.4	.0353	.0370
+.25 to +.30	661-1/2	711.3	.0414	.0445
+.20 to +.25	812-1/2	821.6	.0508	.0514
+.15 to +.20	853-1/2	919.1	.0534	.0575
+.10 to +.15	941	1,000.6	.0589	.0626
+.05 to +.10	1,050-1/2	1,061.3	.0657	.0664
.00 to +.05	1,092-1/2	1,093.3	.0683	.0684
-.05 to .00	1,102-1/2	1,101.3	.0690	.0689
-.10 to -.05	1,103	1,085.3	.0690	.0679
-.15 to -.10	1,086-1/2	1,037.4	.0680	.0649
-.20 to -.15	979	967.0	.0612	.0605
-.25 to -.20	914	875.9	.0572	.0548
-.30 to -.25	820-1/2	772.0	.0513	.0483
-.35 to -.30	704	652.1	.0440	.0408
-.40 to -.35	542	537.1	.0339	.0336
-.45 to -.40	417	417.2	.0261	.0261
-.50 to -.45	320-1/2	310.1	.0201	.0194
-.55 to -.50	186-1/2	217.4	.0117	.0136
-.60 to -.55	123	137.5	.0077	.0086
-.65 to -.60	60-1/2	81.5	.0038	.0051
-.70 to -.65	25	41.6	.0016	.0026
-.75 to -.70	11	17.6	.0007	.0011

Table 46 (continued)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
-.80 to -.75	6	4.8	.0004	.0003
-.85 to -.80	2	1.6	.0001	.0001
Total	15,984	15,984.0	.9999	1.0000

Table 47

Obtained and Theoretical Sampling Distributions of the Product Moment Correlations Coefficient Based on Normal Population (ρ = +.016, N = 30)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
+.60 to +.65	4	1.6	.0001	.0002
+.55 to +.60	5	5.6	.0006	.0007
+.50 to +.55	13-1/2	15.2	.0017	.0019
+.45 to +.50	43-1/2	36.8	.0054	.0046
+.40 to +.45	73-1/2	77.5	.0092	.0097
+.35 to +.40	138-1/2	140.7	.0173	.0176
+.30 to +.35	220	231.0	.0275	.0289
+.25 to +.30	327-1/2	344.5	.0410	.0431
+.20 to +.25	489-1/2	473.9	.0612	.0593
+.15 to +.20	545-1/2	601.8	.0683	.0753
+.10 to +.15	770-1/2	714.5	.0964	.0894
+.05 to +.10	801	792.8	.1002	.0992
.00 to +.05	824-1/2	824.8	.1032	.1032
-.05 to .00	817-1/2	808.0	.1023	.1011
-.10 to -.05	798-1/2	743.3	.0999	.0930
-.15 to -.10	642	641.0	.0803	.0802
-.20 to -.15	501	516.3	.0627	.0646
-.25 to -.20	392	389.2	.0490	.0487
-.30 to -.25	260-1/2	270.9	.0326	.0339
-.35 to -.30	158	173.4	.0198	.0217
-.40 to -.35	81	99.9	.0101	.0125
-.45 to -.40	46	52.7	.0058	.0066
-.50 to -.45	27-1/2	23.2	.0034	.0029
-.55 to -.50	8-1/2	9.6	.0011	.0012
-.60 to -.55	2	3.2	.0003	.0004
-.65 to -.60	1	.8	.0001	.0001
Total	7,992	7,992.2	.9995	1.0000

Table 48

Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Rectangular Population (ρ = -.019, N = 30)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
+.60 to +.65	1	.8	.0001	.0001
+.55 to +.60	1	3.2	.0001	.0004
+.50 to +.55	7	8.8	.0009	.0011
+.45 to +.50	24-1/2	22.4	.0031	.0028
+.40 to +.45	53	50.3	.0066	.0063
+.35 to +.40	83	97.5	.0104	.0122
+.30 to +.35	167-1/2	167.8	.0210	.0210
+.25 to +.30	243-1/2	263.7	.0305	.0330
+.20 to +.25	353-1/2	382.0	.0442	.0478
+.15 to +.20	544	509.1	.0681	.0637
+.10 to +.15	615-1/2	633.8	.0770	.0793
+.05 to +.10	750	737.7	.0938	.0923
.00 to +.05	841	804.8	.1052	.1007
-.05 to .00	843	825.6	.1055	.1033
-.10 to -.05	797	796.0	.0997	.0996
-.15 to -.10	681	720.9	.0852	.0902
-.20 to -.15	628-1/2	609.8	.0786	.0763
-.25 to -.20	466	481.1	.0583	.0602
-.30 to -.25	352-1/2	352.4	.0441	.0441
-.35 to -.30	270-1/2	237.4	.0338	.0297
-.40 to -.35	139	144.7	.0174	.0181
-.45 to -.40	58	79.9	.0073	.0100
-.50 to -.45	48-1/2	39.2	.0061	.0049
-.55 to -.50	15-1/2	16.0	.0019	.0020
-.60 to -.55	4	5.6	.0005	.0007
-.65 to -.60	2	1.6	.0003	.0002
-.70 to -.65	2	.0	.0003	.0000
Total	7,992	7,992.1	1.0000	1.0000

Table 49

Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Leptokurtic Population (ρ = -.004, N = 30)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
+.60 to +.65	0	.8	.0000	.0001
+.55 to +.60	2	4.0	.0003	.0005
+.50 to +.55	6-1/2	11.2	.0008	.0014
+.45 to +.50	30-1/2	28.8	.0038	.0036
+.40 to +.45	57-1/2	59.9	.0072	.0075
+.35 to +.40	99	115.1	.0124	.0144
+.30 to +.35	180	193.4	.0225	.0242
+.25 to +.30	272-1/2	296.5	.0341	.0371
+.20 to +.25	408	420.4	.0511	.0526
+.15 to +.20	570	549.8	.0713	.0688
+.10 to +.15	713-1/2	668.9	.0893	.0837
+.05 to +.10	744	764.0	.0931	.0956
.00 to +.05	851	816.8	.1065	.1022
-.05 to .00	868	820.8	.1086	.1027
-.10 to -.05	790	776.8	.0988	.0972
-.15 to -.10	704-1/2	688.1	.0882	.0861
-.20 to -.15	570	569.8	.0713	.0713
-.25 to -.20	427-1/2	441.2	.0535	.0552
-.30 to -.25	324	315.7	.0405	.0395
-.35 to -.30	173	208.6	.0216	.0261
-.40 to -.35	110-1/2	124.7	.0138	.0156
-.45 to -.40	37	67.1	.0046	.0084
-.50 to -.45	33	31.2	.0041	.0039
-.55 to -.50	18	12.8	.0023	.0016
-.60 to -.55	1	4.0	.0001	.0005
-.65 to -.60	1	1.6	.0001	.0002
Total	7,992	7,992.0	.9999	1.0000

Table 50

Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Slightly Skewed Population ($\rho = -.010$, N = 30)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent Frequency	Theoretical Per Cent Frequency
+.60 to +.65	0	.8	.0000	.0001
+.55 to +.60	5	3.2	.0006	.0004
+.50 to +.55	16	10.4	.0020	.0013
+.45 to +.50	34	26.4	.0043	.0033
+.40 to +.45	63	55.9	.0079	.0070
+.35 to +.40	103-1/2	107.1	.0130	.0134
+.30 to +.35	168	183.8	.0210	.0230
+.25 to +.30	265	282.9	.0332	.0354
+.20 to +.25	369	405.2	.0462	.0507
+.15 to +.20	499-1/2	533.1	.0625	.0667
+.10 to +.15	676	655.3	.0846	.0820
+.05 to +.10	806-1/2	753.6	.1009	.0943
.00 to +.05	823	812.8	.1030	.1017
-.05 to .00	850-1/2	823.2	.1064	.1030
-.10 to -.05	787-1/2	784.8	.0985	.0982
-.15 to -.10	700-1/2	701.7	.0877	.0878
-.20 to -.15	614-1/2	585.8	.0769	.0733
-.25 to -.20	469-1/2	457.2	.0587	.0572
-.30 to -.25	331-1/2	330.1	.0415	.0413
-.35 to -.30	180-1/2	219.8	.0226	.0275
-.40 to -.35	126	132.7	.0158	.0166
-.45 to -.40	52-1/2	71.9	.0066	.0090
-.50 to -.45	29-1/2	34.4	.0037	.0043
-.55 to -.50	16	13.6	.0020	.0017
-.60 to -.55	4	4.8	.0005	.0006
-.65 to -.60	1	1.6	.0001	.0002
Total	7,992	7,992.1	1.0002	1.0000

Table 51

Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Markedly Skewed Population (ρ = -.012, N = 30)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
+.60 to +.65	1	.8	.0001	.0001
+.55 to +.60	6	3.2	.0008	.0004
+.50 to +.55	15	10.4	.0019	.0013
+.45 to +.50	34	25.6	.0043	.0032
+.40 to +.45	53-1/2	54.3	.0067	.0068
+.35 to +.40	131-1/2	104.7	.0165	.0131
+.30 to +.35	173-1/2	179.8	.0217	.0225
+.25 to +.30	282-1/2	279.7	.0353	.0350
+.20 to +.25	385-1/2	398.0	.0482	.0498
+.15 to +.20	544-1/2	529.1	.0681	.0662
+.10 to +.15	622	650.5	.0778	.0814
+.05 to +.10	722-1/2	750.4	.0904	.0939
.00 to +.05	800-1/2	810.4	.1002	.1014
-.05 to .00	790-1/2	824.8	.0989	.1032
-.10 to -.05	806-1/2	787.2	.1009	.0985
-.15 to -.10	747-1/2	705.7	.0935	.0883
-.20 to -.15	594	591.4	.0743	.0740
-.25 to -.20	455	463.5	.0569	.0580
-.30 to -.25	351	334.1	.0439	.0418
-.35 to -.30	244-1/2	223.0	.0306	.0279
-.40 to -.35	126-1/2	135.9	.0158	.0170
-.45 to -.40	64	73.5	.0080	.0092
-.50 to -.45	24-1/2	35.2	.0031	.0044
-.55 to -.50	10-1/2	14.4	.0013	.0018
-.60 to -.55	5-1/2	4.8	.0007	.0006
-.65 to -.60	0	1.6	.0000	.0002
Total	7,992	7,992.0	.9999	1.0000

Table 52

Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Normal Population ($\rho = +.016$, N = 90)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
+.40 to +.45	1	.3	.0004	.0001
+.35 to +.40	0	1.3	.0000	.0005
+.30 to +.35	9	6.7	.0034	.0025
+.25 to +.30	21	25.8	.0079	.0097
+.20 to +.25	66	74.6	.0248	.0280
+.15 to +.20	147-1/2	167.8	.0554	.0630
+.10 to +.15	311	298.4	.1167	.1120
+.05 to +.10	443	425.7	.1663	.1598
.00 to +.05	506	489.4	.1899	.1837
-.05 to .00	453-1/2	457.1	.1702	.1716
-.10 to -.05	371-1/2	346.6	.1395	.1301
-.15 to -.10	185	211.8	.0694	.0795
-.20 to -.15	95	103.4	.0357	.0388
-.25 to -.20	39	40.0	.0146	.0150
-.30 to -.25	10-1/2	12.0	.0039	.0045
-.35 to -.30	4	2.7	.0015	.0010
-.40 to -.35	1	.5	.0004	.0002
Total	2,664	2,664.1	1.0000	1.0000

Table 53

Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Rectangular Population (ρ = −.019, N = 90)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
+.35 to +.40	0	.5	.0000	.0002
+.30 to +.35	0	2.4	.0000	.0009
+.25 to +.30	9	10.9	.0034	.0041
+.20 to +.25	38-1/2	37.6	.0145	.0141
+.15 to +.20	109	98.6	.0409	.0370
+.10 to +.15	190	204.1	.0713	.0766
+.05 to +.10	328	338.3	.1231	.1270
.00 to +.05	448-1/2	452.3	.1684	.1698
−.05 to .00	502	490.2	.1884	.1840
−.10 to −.05	457-1/2	432.1	.1717	.1622
−.15 to −.10	299-1/2	306.6	.1124	.1151
−.20 to −.15	161	174.8	.0604	.0656
−.25 to −.20	81	79.1	.0304	.0297
−.30 to −.25	34	27.4	.0128	.0103
−.35 to −.30	6	7.5	.0023	.0028
−.40 to −.35	0	1.3	.0000	.0005
−.45 to −.40	0	.3	.0000	.0001
Total	2,664	2,664.0	1.0000	1.0000

Table 54

Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Leptokurtic Population (ρ = -.004, N = 90)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequancy
+.35 to +.40	0	.8	.0000	.0003
+.30 to +.35	3	3.7	.0011	.0014
+.25 to +.30	9-1/2	16.3	.0036	.0061
+.20 to +.25	31-1/2	50.9	.0118	.0191
+.15 to +.20	112-1/2	125.2	.0422	.0470
+.10 to +.15	226	243.5	.0848	.0914
+.05 to +.10	390	378.0	.1464	.1419
.00 to +.05	527	473.9	.1978	.1779
-.05 to .00	545-1/2	482.2	.2048	.1810
-.10 to -.05	382-1/2	398.0	.1436	.1494
-.15 to -.10	243-1/2	265.1	.0914	.0995
-.20 to -.15	125	141.5	.0469	.0531
-.25 to -.20	50-1/2	59.7	.0190	.0224
-.30 to -.25	14-1/2	19.4	.0054	.0073
-.35 to -.30	2	4.8	.0008	.0018
-.40 to -.35	1	1.1	.0004	.0004
Total	2,664	2,664.1	1.0000	1.0000

Table 55

Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Slightly Skewed Population (ρ- -.010, N - 90)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
+.35 to +.40	1	.5	.0004	.0002
+.30 to +.35	4	3.2	.0015	.0012
+.25 to +.30	20	14.1	.0075	.0053
+.20 to +.25	42-1/2	45.0	.0160	.0169
+.15 to +.20	105-1/2	114.0	.0396	.0428
+.10 to +.15	230-1/2	227.2	.0865	.0853
+.05 to +.10	356-1/2	362.8	.1338	.1362
.00 to +.05	469-1/2	465.9	.1762	.1749
-.05 to .00	480-1/2	486.7	.1804	.1827
-.10 to -.05	429	412.1	.1610	.1547
-.15 to -.10	304-1/2	281.6	.1143	.1057
-.20 to -.15	139-1/2	154.2	.0524	.0579
-.25 to -.20	55	67.1	.0206	.0252
-.30 to -.25	24	22.4	.0090	.0084
-.35 to -.30	1	5.6	.0004	.0021
-.40 to -.35	1	1.3	.0004	.0005
-.45 to -.40	0	.3	.0000	.0001
Total	2,664	2,664.0	1.0000	1.0001

Table 56

Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Markedly Skewed Population ($\rho = -.012$, N = 90)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
+.35 to +.40	0	.5	.0000	.0002
+.30 to +.35	5	3.2	.0019	.0012
+.25 to +.30	16-1/2	13.1	.0062	.0049
+.20 to +.25	39	43.4	.0146	.0163
+.15 to +.20	107	110.6	.0402	.0415
+.10 to +.15	217	222.2	.0815	.0834
+.05 to +.10	358	357.0	.1344	.1340
.00 to +.05	494-1/2	463.5	.1856	.1740
-.05 to .00	471-1/2	487.5	.1770	.1830
-.10 to -.05	420	416.6	.1577	.1564
-.15 to -.10	271	287.2	.1017	.1078
-.20 to -.15	170	159.0	.0638	.0597
-.25 to -.20	70-1/2	69.3	.0265	.0260
-.30 to -.25	20	23.4	.0075	.0088
-.35 to -.30	4	6.1	.0015	.0023
-.40 to -.35	0	1.1	.0000	.0004
-.45 to -.40	0	.3	.0000	.0001
Total	2,664	2,664.0	1.0001	1.0000

Table 57

Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Normal Population ($\rho = +.835$, N = 15)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
+.975 to +1.000	15-1/2	4.8	.0010	.0003
+.950 to + .975	280	233.4	.0175	.0146
+.925 to + .950	1,048	939.9	.0656	.0588
+.900 to + .925	1,781	1,648.0	.1114	.1031
+.875 to + .900	2,155-1/2	2,004.4	.1349	.1254
+.850 to + .875	2,153-1/2	2,023.6	.1347	.1266
+.825 to + .850	1,898	1,846.2	.1187	.1155
+.800 to + .825	1,564-1/2	1,577.6	.0979	.0987
+.775 to + .800	1,246	1,294.7	.0780	.0810
+.750 to + .775	1,005-1/2	1,032.6	.0629	.0646
+.725 to + .750	714-1/2	812.0	.0447	.0508
+.700 to + .725	549	626.6	.0343	.0392
+.675 to + .700	370	479.5	.0231	.0300
+.650 to + .675	297	364.4	.0186	.0228
+.625 to + .650	248-1/2	276.5	.0155	.0173
+.600 to + .625	172	207.8	.0108	.0130
+.550 to + .600	223	273.3	.0140	.0171
+.500 to + .550	113-1/2	153.4	.0071	.0096
+.450 to + .500	70	84.7	.0044	.0053
+.400 to + .450	36-1/2	46.4	.0023	.0029
+.350 to + .400	20	25.6	.0013	.0016
+.300 to + .350	7-1/2	12.8	.0005	.0008
+.250 to + .300	6	8.0	.0004	.0005
+.200 to + .250	6	3.2	.0004	.0002
+.150 to + .200	0	3.2	.0000	.0002
+.100 to + .150	2	.8	.0001	.0000
+.050 to + .100	0	.8	.0000	.0000
.000 to + .050	1	.0	.0001	.0000
Total	15,984	15,984.2	1,0002	.9999

Table 58

Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Rectangular Population ($\rho = +.831$, N = 15)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
+.975 to +1.000	3-1/2	4.8	.0002	.0003
+.950 to + .975	98-1/2	207.8	.0062	.0130
+.925 to + .950	552	867.9	.0345	.0543
+.900 to + .925	1,346	1,563.2	.0842	.0978
+.875 to + .900	2,184-1/2	1,937.3	.1367	.1212
+.850 to + .875	2,581	1,991.6	.1615	.1246
+.825 to + .850	2,484	1,841.4	.1554	.1152
+.800 to + .825	2,045-1/2	1,593.6	.1280	.0997
+.775 to + .800	1,501-1/2	1,321.9	.0939	.0827
+.750 to + .775	1,089	1,064.5	.0681	.0666
+.725 to + .750	726-1/2	842.4	.0455	.0527
+.700 to + .725	475-1/2	658.5	.0297	.0412
+.675 to + .700	321	506.7	.0201	.0317
+.650 to + .675	217	386.8	.0136	.0242
+.625 to + .650	129-1/2	295.7	.0081	.0185
+.600 to + .625	72-1/2	225.4	.0045	.0141
+.550 to + .600	75-1/2	297.3	.0047	.0186
+.500 to + .550	35-1/2	167.8	.0022	.0105
+.450 to + .500	26-1/2	94.3	.0017	.0059
+.400 to + .450	12	52.7	.0008	.0033
+.350 to + .400	2	28.8	.0001	.0018
+.300 to + .350	3	16.0	.0002	.0010
+.250 to + .300	0	8.0	.0000	.0005
+.200 to + .250	1	4.8	.0001	.0003
+.150 to + .200	1	1.6	.0001	.0001
+.100 to + .150	0	1.6	.0000	.0001
+.050 to + .100	0	1.6	.0000	.0001
.000 to + .050	0	.0	.0000	.0000
Total	15,984	15,984.0	1.0001	1.0000

Table 59

Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Leptokurtic Population (ρ = +.809, N = 15)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
+.975 to +1.000	36	1.6	.0023	.0001
+.950 to + .975	432-1/2	111.9	.0271	.0070
+.925 to + .950	992-1/2	559.4	.0621	.0350
+.900 to + .925	1,400-1/2	1,147.7	.0876	.0718
+.875 to + .900	1,634-1/2	1,582.4	.1023	.0990
+.850 to + .875	1,620	1,769.4	.1014	.1107
+.825 to + .850	1,541-1/2	1,759.8	.0964	.1101
+.800 to + .825	1,396	1,622.4	.0873	.1015
+.775 to + .800	1,216-1/2	1,422.6	.0761	.0890
+.750 to + .775	1,052	1,206.8	.0658	.0755
+.725 to + .750	850	1,000.6	.0532	.0626
+.700 to + .725	733-1/2	813.6	.0459	.0509
+.675 to + .700	587-1/2	653.7	.0368	.0409
+.650 to + .675	457-1/2	517.9	.0286	.0324
+.625 to + .650	368-1/2	409.2	.0231	.0256
+.600 to + .625	316-1/2	322.9	.0198	.0202
+.550 to +..600	460-1/2	444.4	.0288	.0278
+.500 to + .550	297-1/2	266.9	.0186	.0167
+.450 to + .500	172	156.6	.0108	.0098
+.400 to + .450	155	92.7	.0097	.0058
+.350 to + .400	93	52.7	.0058	.0033
+.300 to + .350	59-1/2	30.4	.0037	.0019
+.250 to + .300	49	17.6	.0031	.0011
+.200 to + .250	23	9.6	.0014	.0006
+.150 to + .200	16	4.8	.0010	.0003
+.100 to + .150	9	3.2	.0006	.0002
+.050 to + .100	7	1.6	.0004	.0001
.000 to + .050	5	1.6	.0003	.0001
-.050 to - .000	1	.0	.0001	.0000
-.100 to - .050	0	.0	.0000	.0000
-.150 to - .100	0	.0	.0000	.0000
-.200 to - .150	0	.0	.0000	.0000

Table 59 (continued)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
-.250 to - .200	0	.0	.0000	.0000
-.300 to - .250	1	.0	.0001	.0000
Total	15,984	15,984.0	1.0002	1.0000

Table 60

Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Slightly Skewed Population (ρ = +.826, N = 15)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
+.975 to +1.000	20-1/2	3.2	.0013	.0002
+.950 to + .975	316-1/2	180.6	.0198	.0113
+.925 to + .950	1,039	786.4	.0650	.0492
+.900 to + .925	1,638-1/2	1,459.3	.1025	.0913
+.875 to + .900	1,982-1/2	1,855.7	.1240	.1161
+.850 to + .875	1,963-1/2	1,945.3	.1228	.1217
+.825 to + .850	1,811	1,830.2	.1133	.1145
+.800 to + .825	1,542-1/2	1,608.0	.0965	.1006
+.775 to + .800	1,224	1,352.2	.0766	.0846
+.750 to + .775	1,024-1/2	1,101.3	.0641	.0689
+.725 to + .750	789-1/2	882.3	.0494	.0552
+.700 to + .725	574-1/2	693.7	.0359	.0434
+.675 to + .700	483	540.3	.0302	.0338
+.650 to + .675	365	420.4	.0228	.0263
+.625 to + .650	289-1/2	321.3	.0179	.0201
+.600 to + .625	204-1/2	246.2	.0128	.0154
+.550 to + .600	285-1/2	329.3	.0179	.0206
+.500 to + .550	166-1/2	188.6	.0104	.0118
+.450 to + .500	105-1/2	105.5	.0066	.0066
+.400 to + .450	64-1/2	60.7	.0040	.0038
+.350 to + .400	31-1/2	33.6	.0020	.0021
+.300 to + .350	29	19.2	.0018	.0012
+.250 to + .300	20-1/2	9.6	.0013	.0006
+.200 to + .250	4-1/2	4.8	.0003	.0003
+.150 to + .200	5	3.2	.0003	.0002
+.100 to + .150	3	1.6	.0002	.0001
+.050 to + .100	1	1.6	.0001	.0001
.000 to + .050	1	0.0	.0001	.0000
-.050 to .000	1	0.0	.0001	.0000
Total	15,984	15,984.1	1.0000	1.0000

Table 61

Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Markedly Skewed Population ($\rho = +.826$, N = 15)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
+.975 to +1.000	21-1/2	3.2	.0013	.0002
+.950 to + .975	420-1/2	180.6	.0263	.0113
+.925 to + .950	1,142-1/2	786.4	.0715	.0492
+.900 to + .925	1,716	1,459.3	.1074	.0913
+.875 to + .900	1,975	1,855.7	.1236	.1161
+.850 to + .875	1,858	1,945.3	.1162	.1217
+.825 to + .850	1,661-1/2	1,830.2	.1039	.1145
+.800 to + .825	1,419	1,608.0	.0888	.1006
+.775 to + .800	1,155	1,352.2	.0723	.0846
+.750 to + .775	987-1/2	1,101.3	.0618	.0689
+.725 to + .750	792	882.3	.0495	.0552
+.700 to + .725	630-1/2	693.7	.0394	.0434
+.675 to + .700	484	540.3	.0303	.0338
+.650 to + .675	384	420.4	.0240	.0263
+.625 to + .650	291	321.3	.0182	.0201
+.600 to + .625	221-1/2	246.2	.0139	.0154
+.550 to + .600	331-1/2	329.3	.0207	.0206
+.500 to + .550	170-1/2	188.6	.0107	.0118
+.450 to + .500	130-1/2	105.5	.0082	.0066
+.400 to + .450	80	60.7	.0050	.0038
+.350 to + .400	43	33.6	.0027	.0021
+.300 to + .350	25-1/2	19.2	.0016	.0012
+.250 to + .300	14-1/2	9.6	.0009	.0006
+.200 to + .250	10	4.8	.0006	.0003
+.150 to + .200	6	3.2	.0004	.0002
+.100 to + .150	4	1.6	.0003	.0001
+.050 to + .100	7	1.6	.0004	.0001
.000 to + .050	2	0.0	.0001	.0000
Total	15,984	15,984.1	1.0000	1.0000

Table 62

Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Normal Population (ρ = +.835, N = 30)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
+.950 to +.975	7-1/2	4.8	.0009	.0006
+.925 to +.950	138-1/2	114.3	.0173	.0143
+.900 to +.925	527	536.3	.0659	.0671
+.875 to +.900	1,182-1/2	1,090.9	.1480	.1365
+.850 to +.875	1,462-1/2	1,401.8	.1830	.1754
+.825 to +.850	1,436	1,377.0	.1797	.1723
+.800 to +.825	1,068	1,139.7	.1336	.1426
+.775 to +.800	787	840.0	.0985	.1051
+.750 to +.775	559-1/2	570.6	.0700	.0714
+.725 to +.750	333-1/2	370.0	.0417	.0463
+.700 to +.725	214-1/2	227.8	.0268	.0285
+.675 to +.700	127-1/2	135.9	.0160	.0170
+.650 to +.675	59	79.9	.0074	.0100
+.625 to +.650	48	45.6	.0060	.0057
+.600 to +.625	13	26.4	.0016	.0033
+.550 to +.600	19	21.6	.0024	.0027
+.500 to +.550	7	7.2	.0009	.0009
+.450 to +.500	2	1.6	.0003	.0002
+.400 to +.450	0	.8	.0000	.0001
Total	7,992	7,992.2	1.0000	1.0000

Table 63

Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Rectangular Population (ρ = +.831, N = 30)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
+.950 to +.975	0	3.2	.0000	.0004
+.925 to +.950	23-1/2	96.7	.0029	.0121
+.900 to +.925	262-1/2	477.9	.0328	.0598
+.875 to +.900	936-1/2	1,013.4	.1172	.1268
+.850 to +.875	1,663	1,351.4	.2081	.1691
+.825 to +.850	1,736	1,369.0	.2172	.1713
+.800 to +.825	1,477	1,164.4	.1848	.1457
+.775 to +.800	938	879.1	.1174	.1100
+.750 to +.775	520	612.2	.0651	.0766
+.725 to +.750	250	402.8	.0313	.0504
+.700 to +.725	111-1/2	253.3	.0140	.0317
+.675 to +.700	37-1/2	154.2	.0047	.0193
+.650 to +.675	15-1/2	91.9	.0019	.0115
+.625 to +.650	12	53.5	.0015	.0067
+.600 to +.625	4	30.4	.0005	.0038
+.550 to +.600	5	26.4	.0006	.0033
+.500 to +.550	0	8.8	.0000	.0011
+.450 to +.500	0	2.4	.0000	.0003
+.400 to +.450	0	.8	.0000	.0001
Total	7,992	7,991.8	1.0000	1.0000

Table 64

Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Leptokurtic Population (ρ = +.809, N = 30)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
+.975 to +1.000	1	.0	.0001	.0000
+.950 to + .975	23	.8	.0029	.0001
+.925 to + .950	200	37.6	.0250	.0047
+.900 to + .925	494	243.0	.0618	.0304
+.875 to + .900	816-1/2	645.8	.1022	.0808
+.850 to + .875	1,026-1/2	1,040.6	.1284	.1302
+.825 to + .850	1,054-1/2	1,238.0	.1319	.1549
+.800 to + .825	951	1,211.6	.1190	.1516
+.775 to + .800	816-1/2	1,038.2	.1022	.1299
+.750 to + .775	674-1/2	808.8	.0844	.1012
+.725 to + .750	484	590.6	.0606	.0739
+.700 to + .725	395-1/2	409.2	.0495	.0512
+.675 to + .700	285	273.3	.0357	.0342
+.650 to + .675	218	175.0	.0273	.0219
+.625 to + .650	180-1/2	111.1	.0226	.0139
+.600 to + .625	111-1/2	68.7	.0140	.0086
+.550 to + .600	130-1/2	65.5	.0163	.0082
+.500 to + .550	69-1/2	23.2	.0087	.0029
+.450 to + .500	32-1/2	8.0	.0041	.0010
+.400 to + .450	13-1/2	2.4	.0017	.0003
+.350 to + .400	8	.8	.0010	.0001
+.300 to + .350	4	.0	.0005	.0000
+.250 to + .300	1	.0	.0001	.0000
+.200 to + .250	0	.0	.0000	.0000
+.150 to + .200	1	.0	.0001	.0000
Total	7,992	7,992.2	1.0001	1.0000

Table 65

Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Slightly Skewed Population (ρ = +.826, N = 30)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
+.950 to +.975	7	2.4	.0009	.0003
+.925 to +.950	148	78.3	.0185	.0098
+.900 to +.925	541-1/2	410.8	.0678	.0514
+.875 to +.900	995-1/2	920.7	.1246	.1152
+.850 to +.875	1,294	1,283.5	.1619	.1606
+.825 to +.850	1,265	1,351.4	.1583	.1691
+.800 to +.825	1,117-1/2	1,188.4	.1398	.1487
+.775 to +.800	891	924.7	.1115	.1157
+.750 to +.775	611	661.7	.0765	.0828
+.725 to +.750	413-1/2	446.0	.0517	.0558
+.700 to +.725	254	286.9	.0318	.0359
+.675 to +.700	162	178.2	.0203	.0223
+.650 to +.675	105-1/2	108.7	.0132	.0136
+.625 to +.650	58-1/2	63.9	.0073	.0080
+.600 to +.625	49	37.6	.0061	.0047
+.550 to +.600	55	33.6	.0069	.0042
+.500 to +.550	16	10.4	.0020	.0013
+.450 to +.500	6	3.2	.0008	.0004
+.400 to +.450	2	1.6	.0003	.0002
Total	7,992	7,992.0	1.0002	1.0000

Table 66

Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Markedly Skewed Population ($\rho = +.826$, N = 30)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
+.950 to +.975	13	2.4	.0016	.0003
+.925 to +.950	163	78.3	.0204	.0098
+.900 to +.925	585-1/2	410.8	.0733	.0514
+.875 to +.900	1,065	920.7	.1333	.1152
+.850 to +.875	1,260-1/2	1,283.5	.1577	.1606
+.825 to +.850	1,201-1/2	1,351.4	.1503	.1691
+.800 to +.825	1,062-1/2	1,188.4	.1329	.1487
+.775 to +.800	813-1/2	924.7	.1018	.1157
+.750 to +.775	610	661.7	.0763	.0828
+.725 to +.750	397	446.0	.0497	.0558
+.700 to +.725	312	286.9	.0390	.0359
+.675 to +.700	180	178.2	.0225	.0223
+.650 to +.675	123-1/2	108.7	.0155	.0136
+.625 to +.650	77-1/2	63.9	.0097	.0080
+.600 to +.625	48-1/2	37.6	.0061	.0047
+.550 to +.600	47-1/2	33.6	.0059	.0042
+.550 to +.550	24-1/2	10.4	.0031	.0013
+.450 to +.500	6	3.2	.0008	.0004
+.400 to +.450	1	1.6	.0001	.0002
Total	7,992	7,992.0	1.0000	1.0000

Table 67

Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Normal Population (ρ - +.835, N - 90)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
+.92 to +.93	0	.5	.0000	.0002
+.91 to +.92	6	2.9	.0023	.0011
+.90 to +.91	9-1/2	13.1	.0036	.0049
+.89 to +.90	40-1/2	40.0	.0152	.0150
+.88 to +.89	72-1/2	89.8	.0272	.0337
+.87 to +.88	156-1/2	160.4	.0587	.0602
+.86 to +.87	237	235.5	.0890	.0884
+.85 to +.86	289-1/2	295.7	.1087	.1110
+.84 to +.85	344	328.2	.1291	.1232
+.83 to +.84	346	326.9	.1299	.1227
+.82 to +.83	287	296.0	.1077	.1111
+.81 to +.82	265	248.6	.0995	.0933
+.80 to +.81	185	195.3	.0694	.0733
+.79 to +.80	147-1/2	145.5	.0554	.0546
+.78 to +.79	124	102.0	.0465	.0383
+.77 to +.78	59	69.3	.0221	.0260
+.76 to +.77	34-1/2	45.0	.0130	.0169
+.75 to +.76	22-1/2	28.2	.0084	.0106
+.74 to +.75	12	17.6	.0045	.0066
+.73 to +.74	12-1/2	10.1	.0047	.0038
+.72 to +.73	10-1/2	6.1	.0039	.0023
+.71 to +.72	1/2	3.5	.0002	.0013
+.70 to +.71	1-1/2	1.9	.0006	.0007
+.69 to +.70	1	1.1	.0004	.0004
+.68 to +.69	0	.5	.0000	.0002
+.67 to +.68	0	.3	.0000	.0001
+.66 to +.67	0	.1	.0000	.0000
+.65 to +.66	0	.1	.0000	.0000
Total	2,664	2,664.2	1.0000	.9999

Table 68

Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Rectangular Population (ρ = +.831, N = 90)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
+.92 to +.93	0	.3	.0000	.0001
+.91 to +.92	0	2.1	.0000	.0008
+.90 to +.91	1	9.3	.0004	.0035
+.89 to +.90	7	30.4	.0026	.0114
+.88 to +.89	23-1/2	71.7	.0088	.0269
+.87 to +.88	83	134.3	.0312	.0504
+.86 to +.87	174	207.3	.0653	.0778
+.85 to +.86	299	271.5	.1122	.1019
+.84 to +.85	379-1/2	313.6	.1425	.1177
+.83 to +.84	470-1/2	324.5	.1766	.1218
+.82 to +.83	409-1/2	304.2	.1537	.1142
+.81 to +.82	305-1/2	265.1	.1147	.0995
+.80 to +.81	208	214.7	.0781	.0806
+.79 to +.80	139	164.6	.0522	.0618
+.78 to +.79	81-1/2	119.9	.0306	.0450
+.77 to +.78	43	83.6	.0162	.0314
+.76 to +.77	22	55.7	.0083	.0209
+.75 to +.76	10-1/2	36.0	.0039	.0135
+.74 to +.75	4-1/2	22.6	.0017	.0085
+.73 to +.74	2-1/2	13.9	.0009	.0052
+.72 to +.73	1/2	8.0	.0002	.0030
+.71 to +.72	0	4.8	.0000	.0018
+.70 to +.71	0	2.7	.0000	.0010
+.69 to +.70	0	1.6	.0000	.0006
+.68 to +.69	0	.8	.0000	.0003
+.67 to +.68	0	.5	.0000	.0002
+.66 to +.67	0	.3	.0000	.0001
+.65 to +.66	0	.1	.0000	.0000
+.64 to +.65	0	.1	.0000	.0000
Total	2,664	2,664.2	1.001	.9999

Table 69

Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Leptokurtic Population (ρ = +.809, N = 90)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
+.92 to +.93	1	.0	.0004	.0000
+.91 to +.92	7	.3	.0026	.0001
+.90 to +.91	23-1/2	1.3	.0088	.0005
+.89 to +.90	39-1/2	5.6	.0148	.0021
+.88 to +.89	52-1/2	17.8	.0197	.0067
+.87 to +.88	92-1/2	43.2	.0347	.0162
+.86 to +.87	139-1/2	84.4	.0524	.0317
+.85 to +.86	163	137.7	.0612	.0517
+.84 to +.85	189	195.8	.0710	.0735
+.83 to +.84	183-1/2	247.2	.0689	.0928
+.82 to +.83	236	279.5	.0886	.1049
+.81 to +.82	217	290.4	.0815	.1090
+.80 to +.81	223	279.2	.0837	.1048
+.79 to +.80	187	251.7	.0702	.0945
+.78 to +.79	170	213.4	.0638	.0801
+.77 to +.78	128	172.9	.0480	.0649
+.76 to +.77	129	133.5	.0484	.0501
+.75 to +.76	109	98.6	.0409	.0370
+.74 to +.75	85-1/2	71.1	.0321	.0267
+.73 to +.74	63-1/2	49.3	.0238	.0185
+.72 to +.73	48	33.3	.0180	.0125
+.71 to +.72	51-1/2	21.6	.0193	.0081
+.70 to +.71	27	14.1	.0101	.0053
+.69 to +.70	29-1/2	8.8	.0111	.0033
+.68 to +.69	15-1/2	5.3	.0058	.0020
+.67 to +.68	15-1/2	3.5	.0058	.0013
+.66 to +.67	12-1/2	1.9	.0047	.0007
+.65 to +.66	7-1/2	1.1	.0028	.0004
+.64 to +.65	3-1/2	.8	.0013	.0003
+.63 to +.64	2-1/2	.3	.0009	.0001
+.62 to +.63	1	.3	.0004	.0001
+.61 to +.62	4	.1	.0015	.0000

Table 69 (continued)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
+.60 to +.61	3	.1	.0011	.0000
+.59 to +.60	1	.0	.0004	.0000
+.58 to +.59	2	.0	.0008	.0000
+.49 to +.50	1	.0	.0004	.0000
Total	2,664	2,664.1	.9999	.9999

Table 70

Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Slightly Skewed Population ($\rho = +.826$, N = 90)

r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
+.92 to +.93	0	.3	.0000	.0001
+.91 to +.92	3	1.1	.0011	.0004
+.90 to +.91	14-1/2	6.1	.0054	.0023
+.89 to +.90	29-1/2	21.3	.0111	.0080
+.88 to +.89	74-1/2	53.3	.0280	.0200
+.87 to +.88	136-1/2	106.0	.0512	.0398
+.86 to +.87	179	173.4	.0672	.0651
+.85 to +.86	275-1/2	239.2	.1034	.0898
+.84 to +.85	266-1/2	290.6	.1000	.1091
+.83 to +.84	294-1/2	314.9	.1105	.1182
+.82 to +.83	288	309.3	.1081	.1161
+.81 to +.82	253-1/2	280.3	.0952	.1052
+.80 to +.81	214	237.1	.0803	.0890
+.79 to +.80	176-1/2	188.9	.0663	.0709
+.78 to +.79	144-1/2	142.3	.0542	.0534
+.77 to +.78	88-1/2	103.1	.0332	.0387
+.76 to +.77	76-1/2	70.6	.0287	.0265
+.75 to +.76	57	47.7	.0214	.0179
+.74 to +.75	22-1/2	30.6	.0084	.0115
+.73 to +.74	18-1/2	19.4	.0069	.0073
+.72 to +.73	9	12.0	.0034	.0045
+.71 to +.72	13	6.9	.0049	.0026
+.70 to +.71	13	4.3	.0049	.0016
+.69 to +.70	4-1/2	2.4	.0017	.0009
+.68 to +.69	3-1/2	1.3	.0013	.0005
+.67 to +.68	3-1/2	.8	.0013	.0003
+.66 to +.67	2-1/2	.3	.0009	.0001
+.65 to +.66	1	.3	.0004	.0001
+.64 to +.65	0	.3	.0000	.0001
+.63 to +.64	1	.0	.0004	.0000
Total	2,664	2,664.1	.9998	1.0000

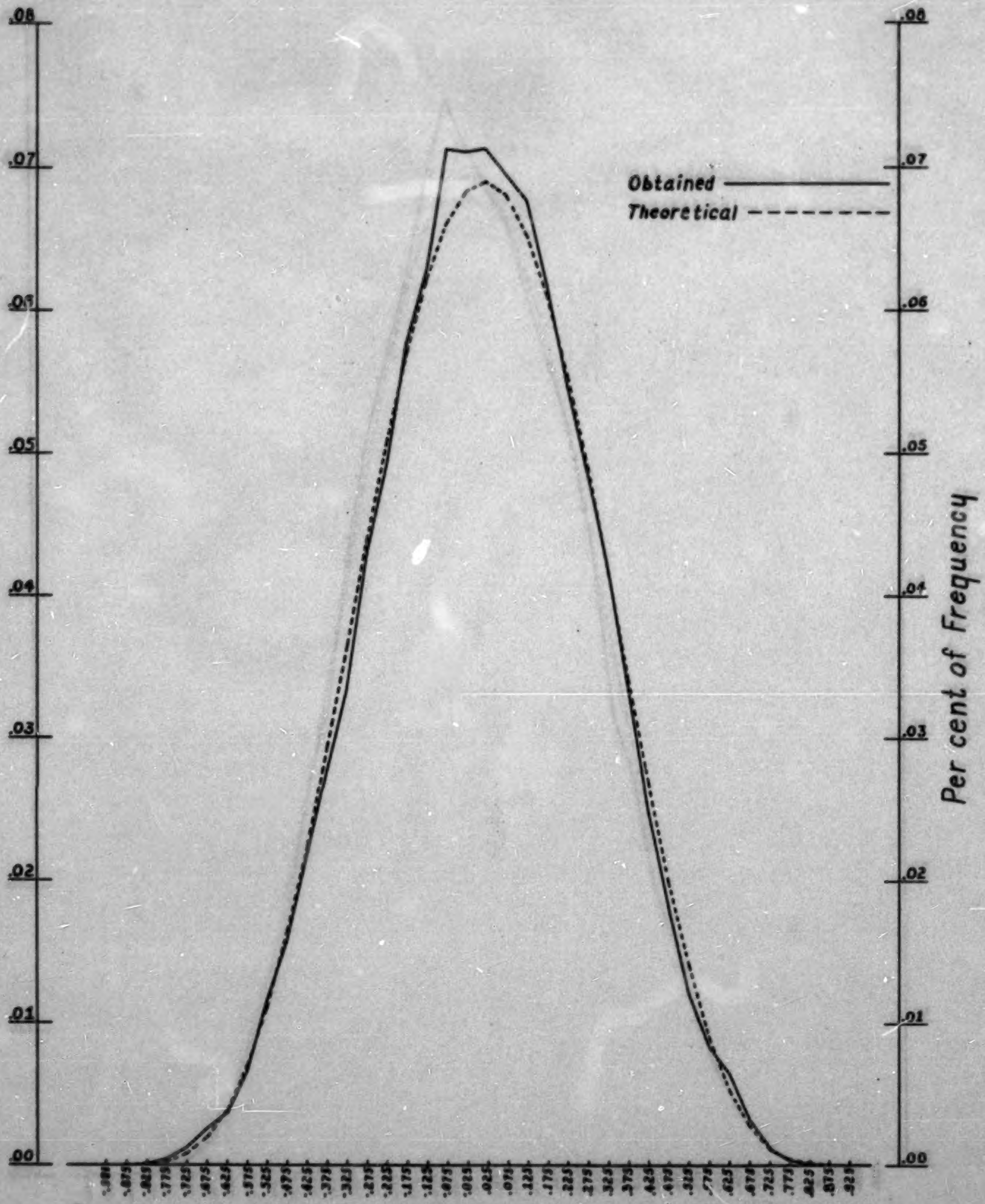
Table 71

Obtained and Theoretical Sampling Distributions of the Product Moment Correlation Coefficient Based on Markedly Skewed Population (ρ = +.826, N = 90)

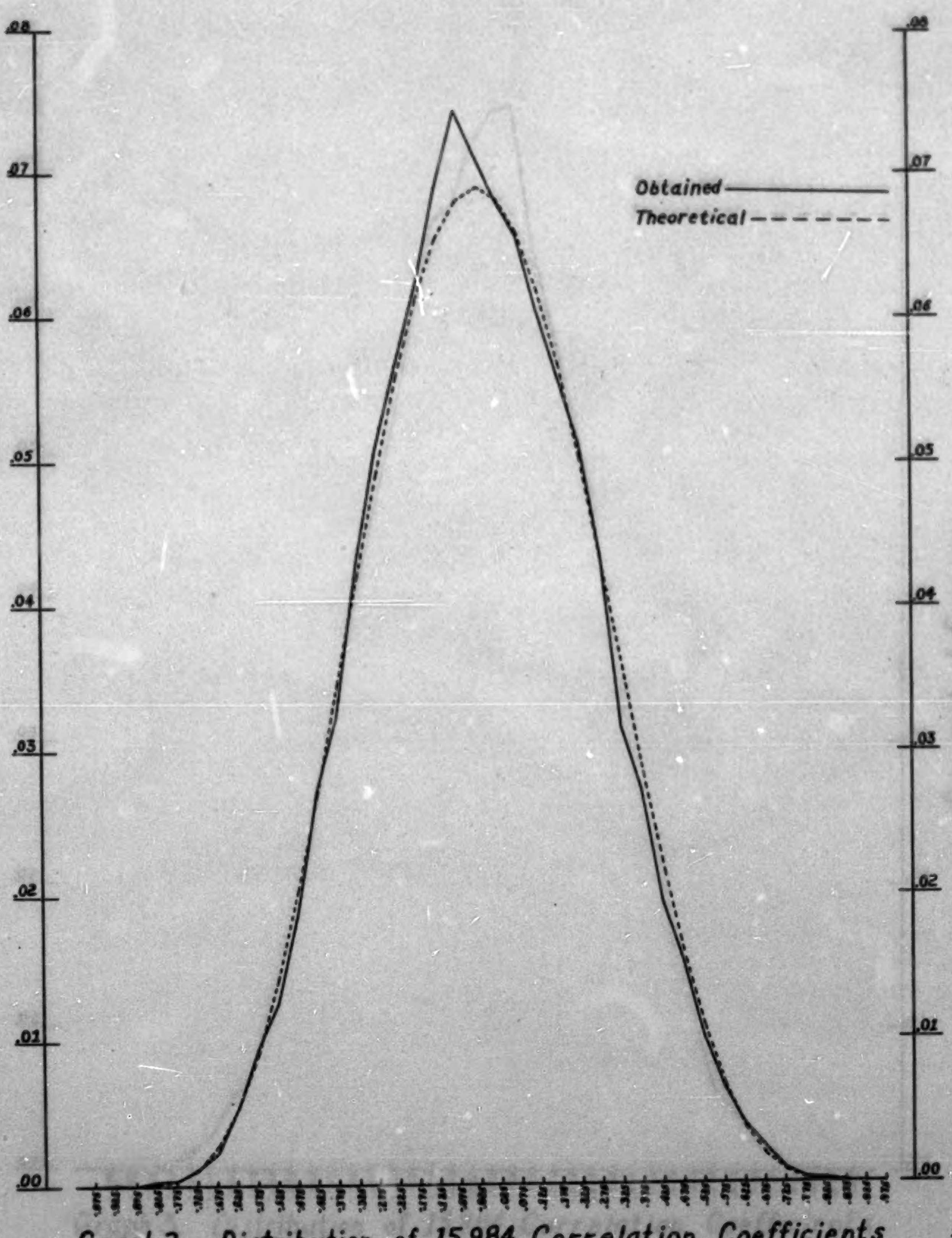
r	Obtained Frequency	Theoretical Frequency	Obtained Per Cent of Frequency	Theoretical Per Cent of Frequency
+.92 to +.93	1	.3	.0004	.0001
+.91 to +.92	5-1/2	1.1	.0021	.0004
+.90 to +.91	18-1/2	6.1	.0069	.0023
+.89 to +.90	42-1/2	21.3	.0160	.0080
+.88 to +.89	77	53.3	.0289	.0200
+.87 to +.88	139-1/2	106.0	.0524	.0398
+.86 to +.87	178-1/2	173.4	.0670	.0651
+.85 to +.86	258-1/2	239.2	.0970	.0898
+.84 to +.85	265	290.6	.0995	.1091
+.83 to +.84	259	314.9	.0972	.1182
+.82 to +.83	278-1/2	309.3	.1045	.1161
+.81 to +.82	261	280.3	.0980	.1052
+.80 to +.81	216	237.1	.0811	.0890
+.79 to +.80	169-1/2	188.9	.0636	.0709
+.78 to +.79	157-1/2	142.3	.0591	.0534
+.77 to +.78	96	103.1	.0360	.0387
+.76 to +.77	78-1/2	70.6	.0295	.0265
+.75 to +.76	51-1/2	47.7	.0193	.0179
+.74 to +.75	47	30.6	.0176	.0115
+.73 to +.74	20	19.4	.0075	.0073
+.72 to +.73	19-1/2	12.0	.0073	.0045
+.71 to +.72	12	6.9	.0045	.0026
+.70 to +.71	4	4.3	.0015	.0016
+.69 to +.70	2	2.4	.0008	.0009
+.68 to +.69	4	1.3	.0015	.0005
+.67 to +.68	1	.8	.0004	.0003
+.66 to +.67	1	.3	.0004	.0001
+.65 to +.66	0	.3	.0000	.0001
+.64 to +.65	0	.3	.0000	.0001
Total	2,664	2,664.1	1.0000	1.0000

APPENDIX C

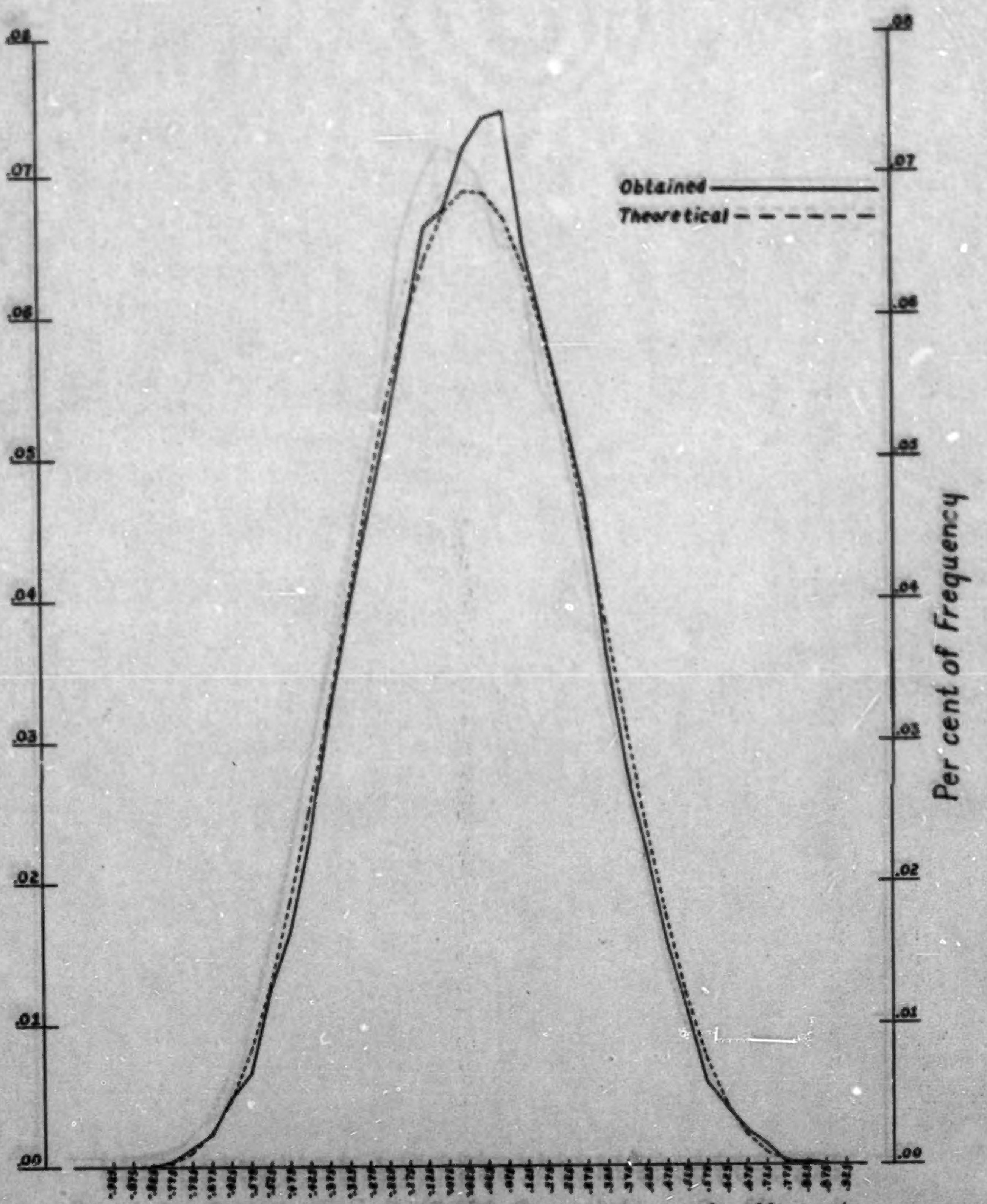
GRAPHS OF SAMPLING DISTRIBUTIONS OF THE PRODUCT MOMENT CORRELATION COEFFICIENT



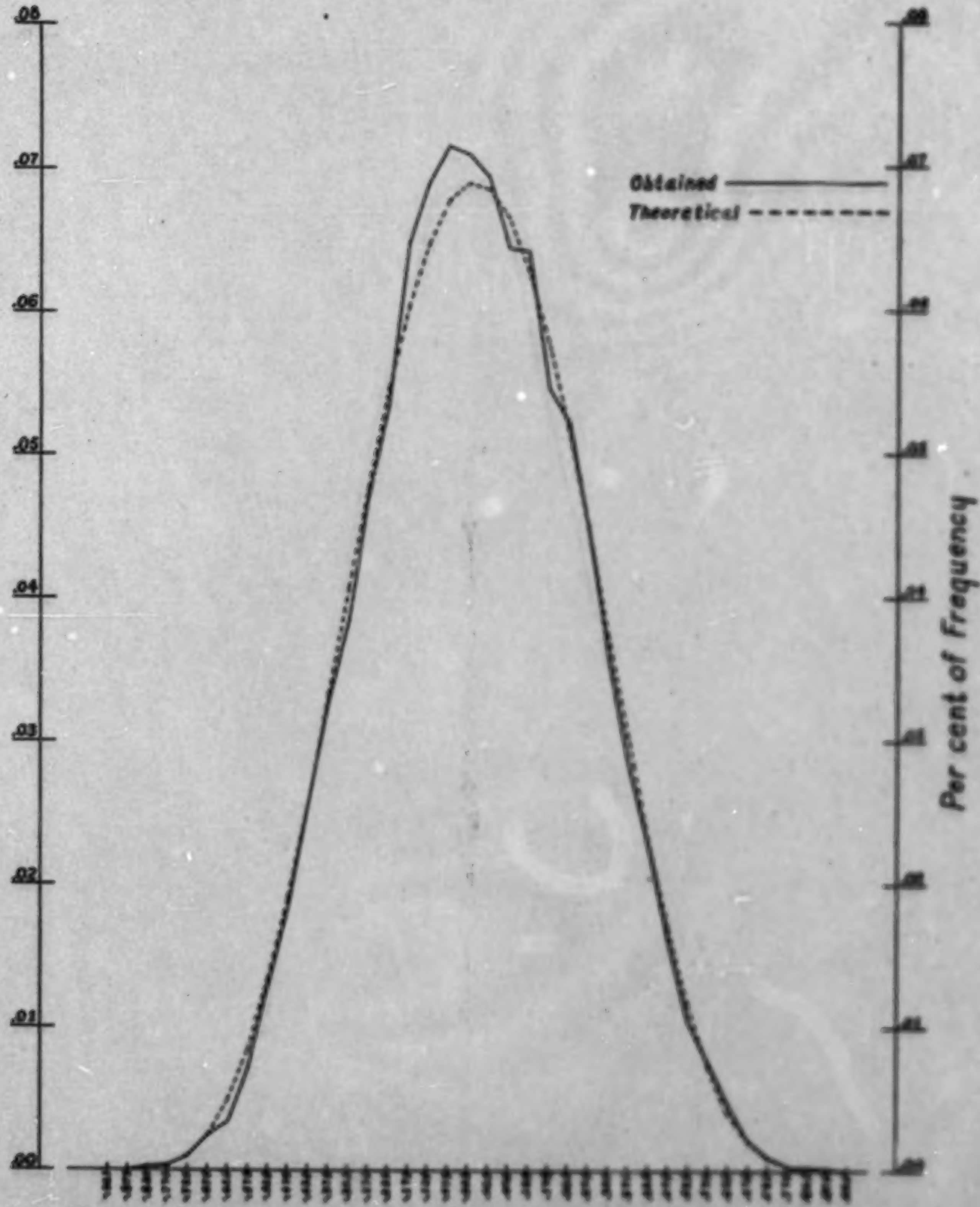
Graph 1. Distribution of 15,984 Correlation Coefficients Based on Normal Population ($\rho = +.016$, $N = 15$)



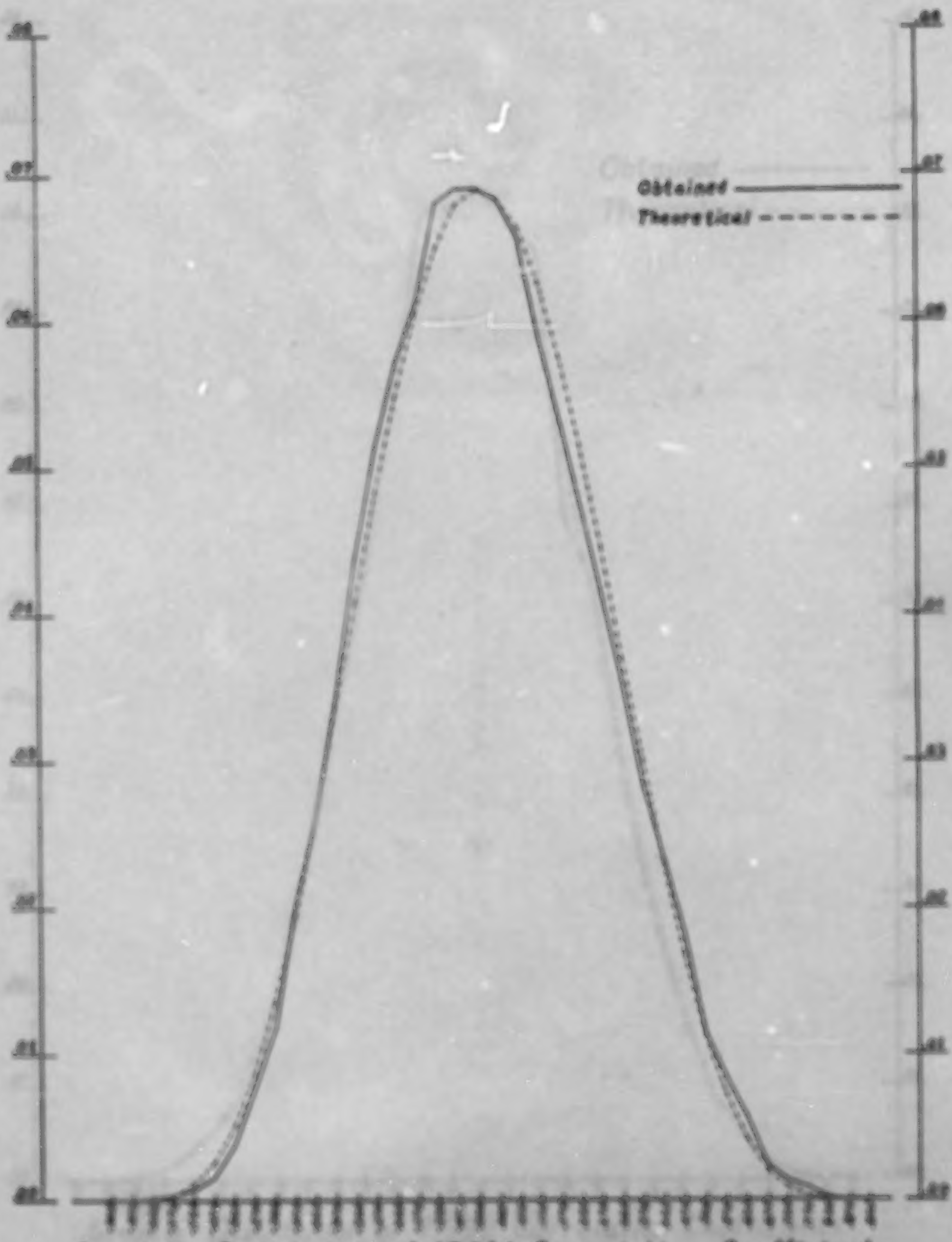
Graph 2. Distribution of 15,984 Correlation Coefficients Based on Rectangular Population ($\rho = -.019$, $N = 15$)



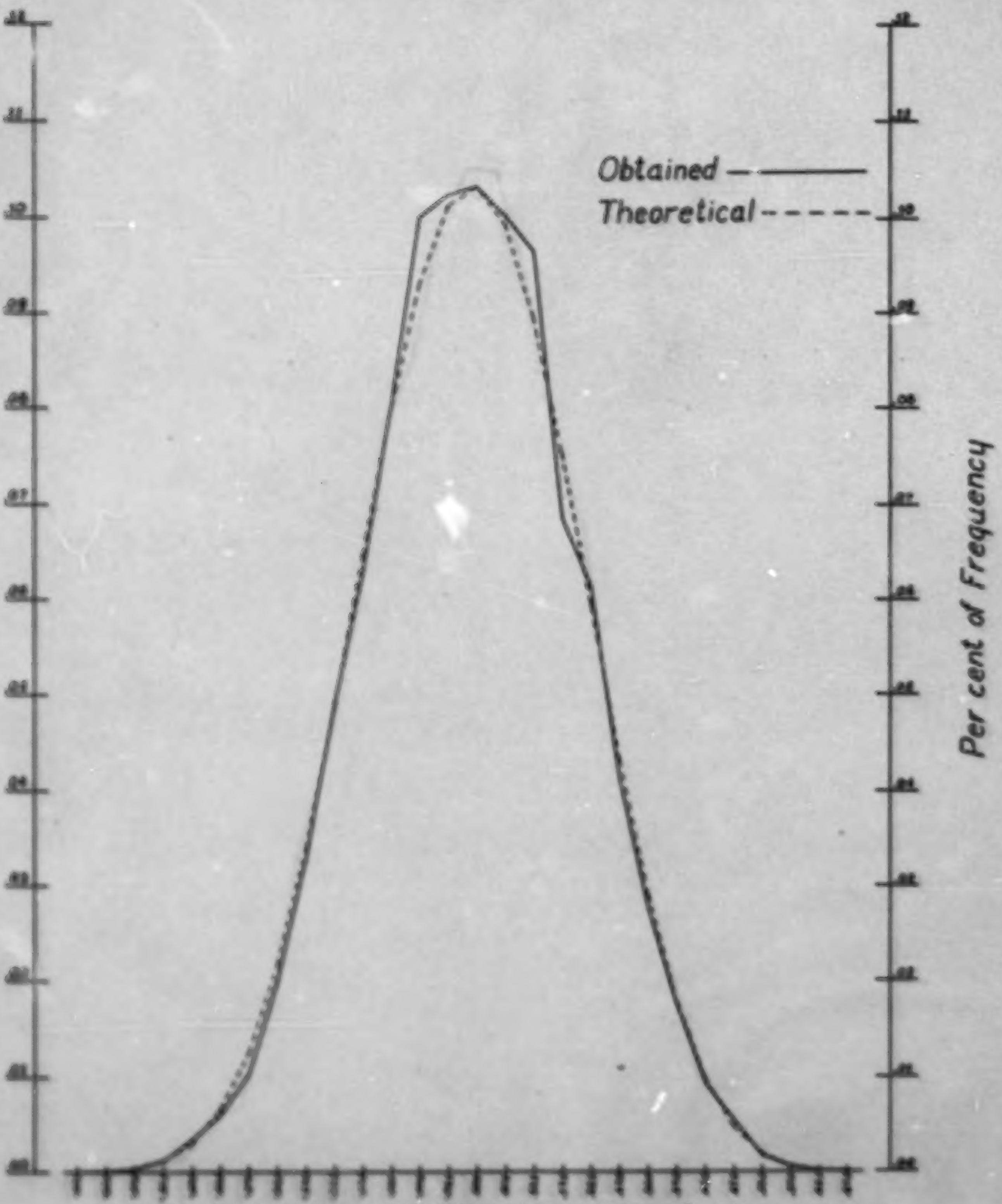
Graph 3. Distribution of 15,984 Correlation Coefficients Based on Leptokurtic Population ($\rho = -.004, N = 15$)



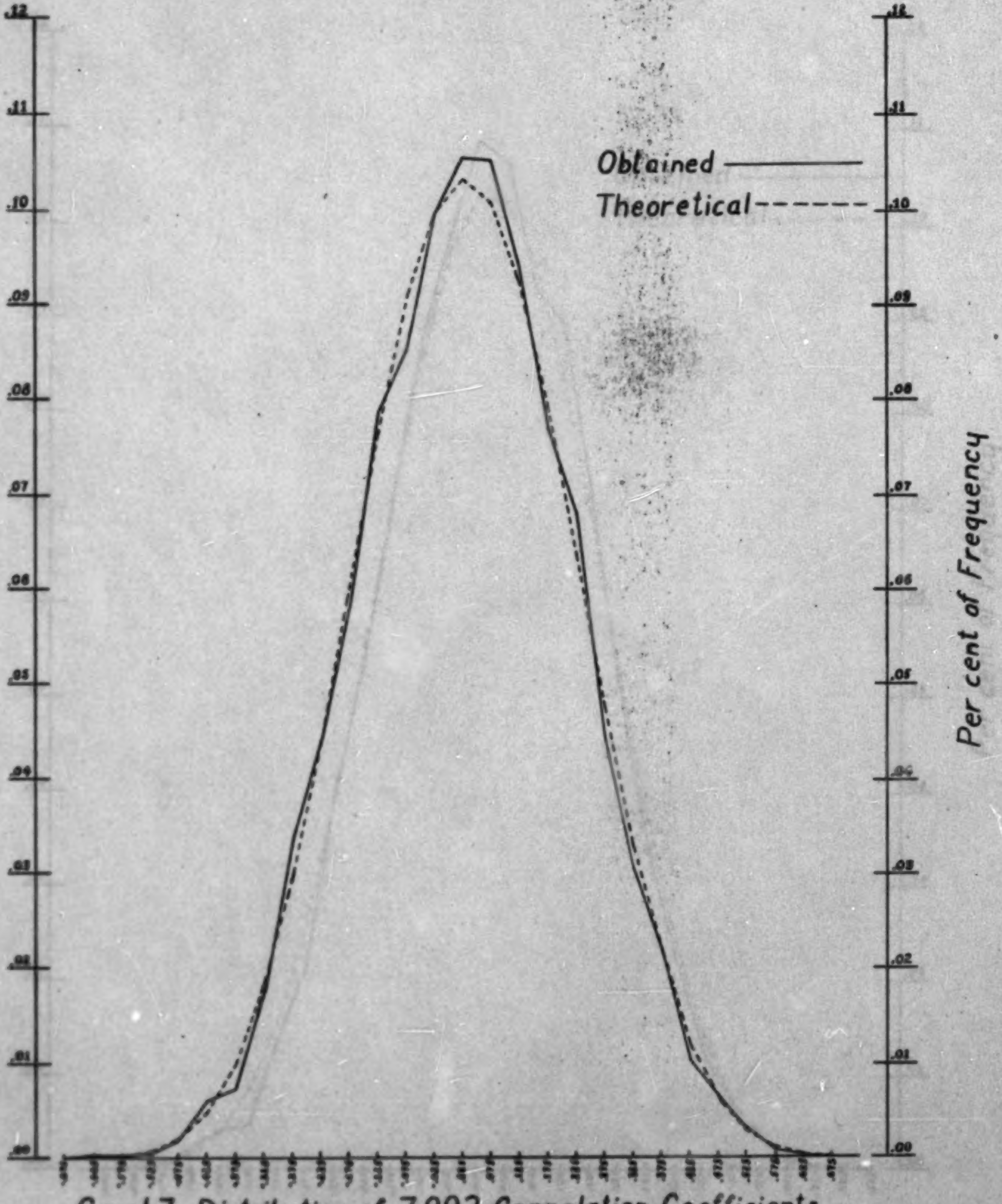
Graph 4. Distribution of 15,984 Correlation Coefficients Based on Slightly Skewed Population ($\rho = -.010$, $N = 15$)



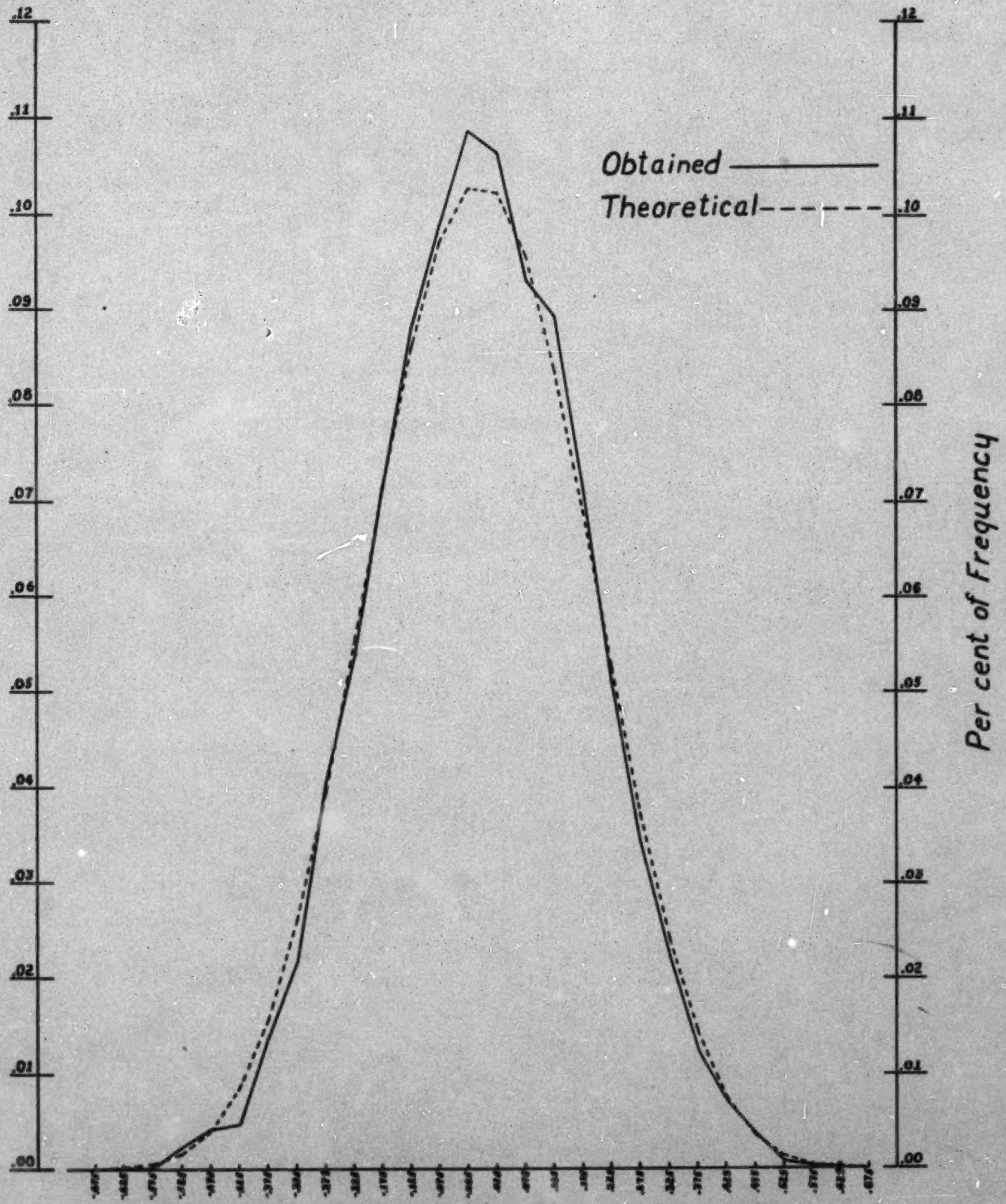
Graph 5. Distribution of 15,984 Correlation Coefficients Based on Markedly Skewed Population ($\rho = -.012$, $N = 15$)



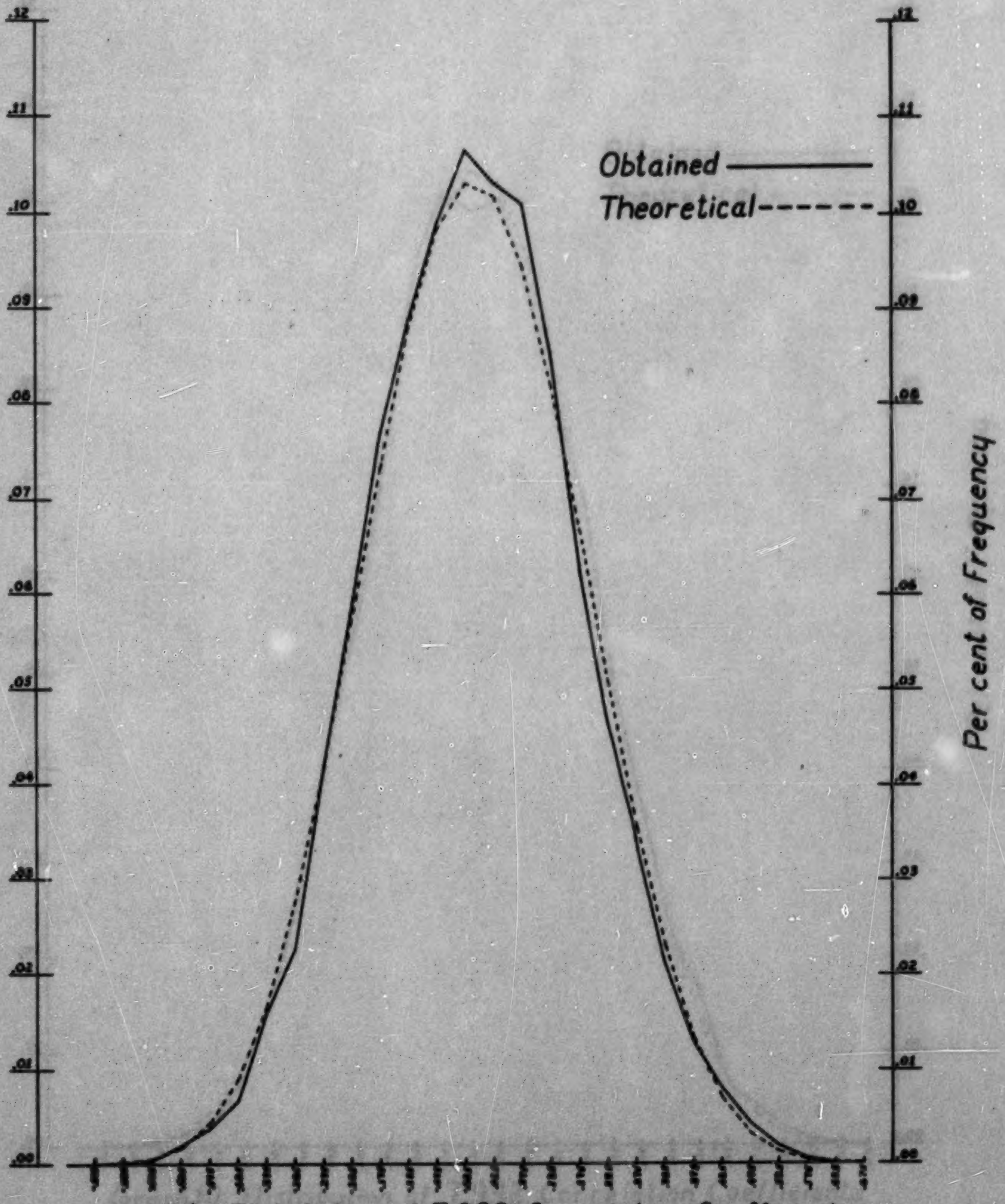
Graph 6. Distribution of 7,992 Correlation Coefficients Based on Normal Population (p = -.016, N = 30)



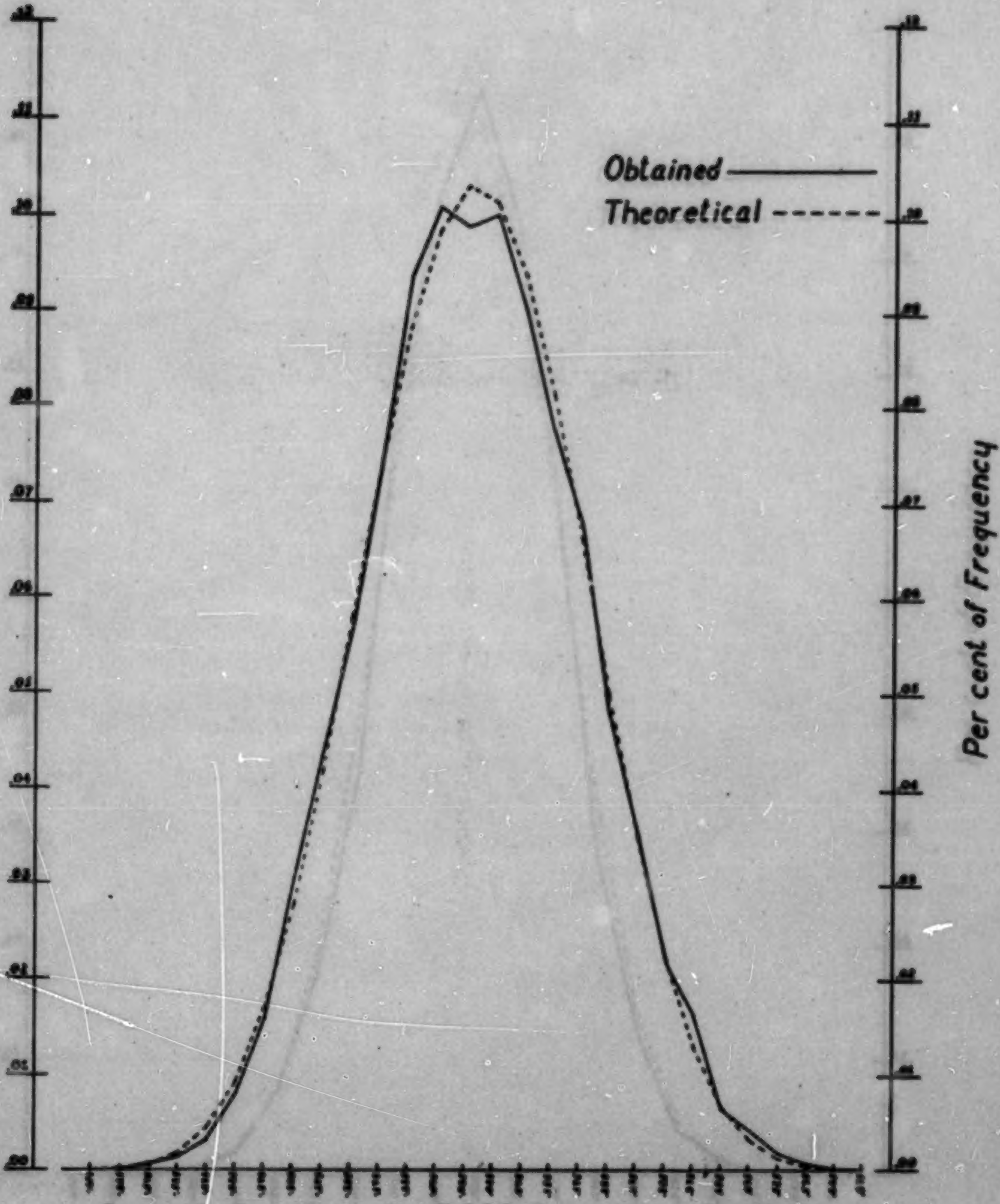
Graph 7. Distribution of 7,992 Correlation Coefficients Based on Rectangular Population (p=-.019, N=30)



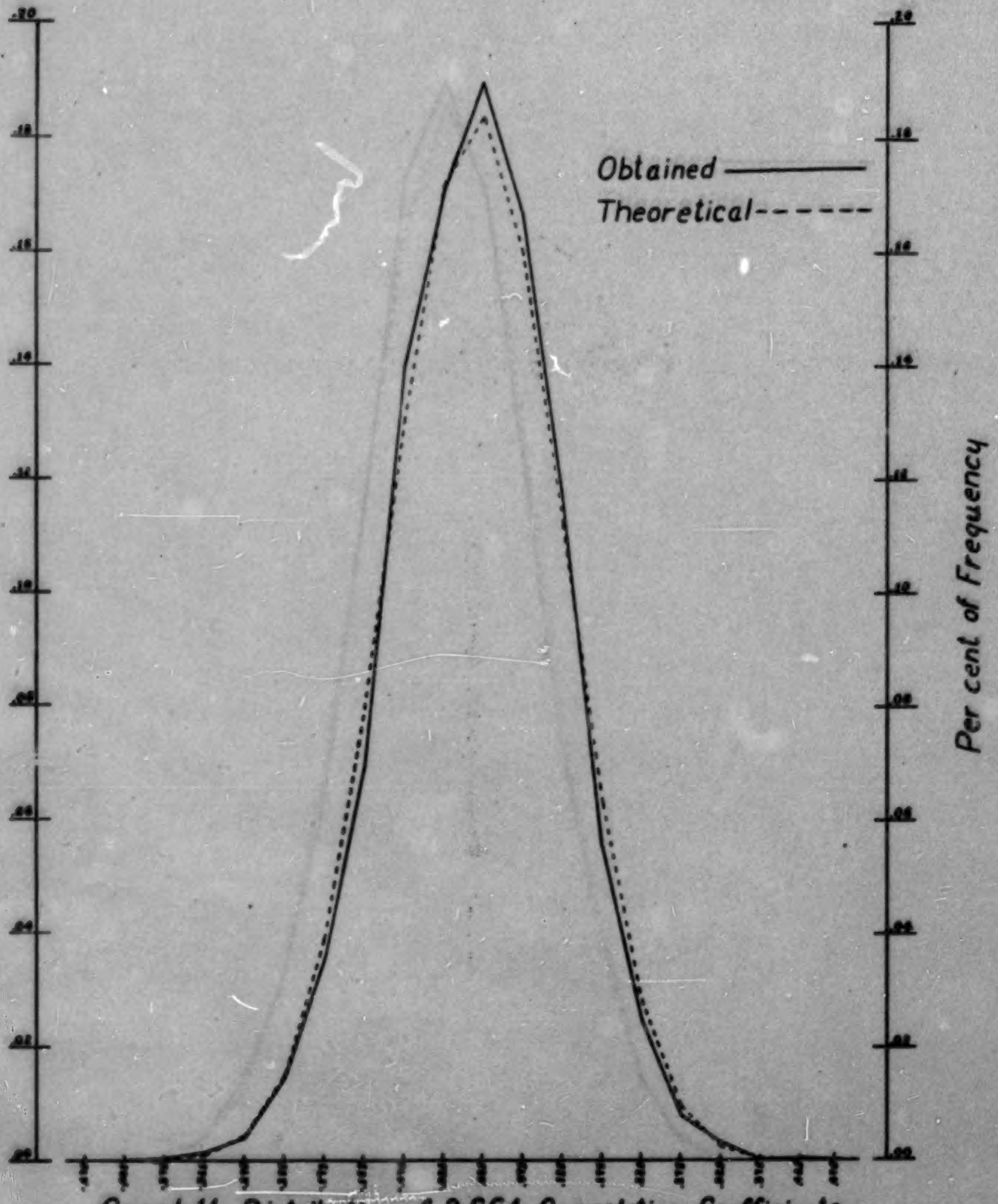
Graph 8. Distribution of 7,992 Correlation Coefficients Based on Leptokurtic Population (p = -.004, N = 30)



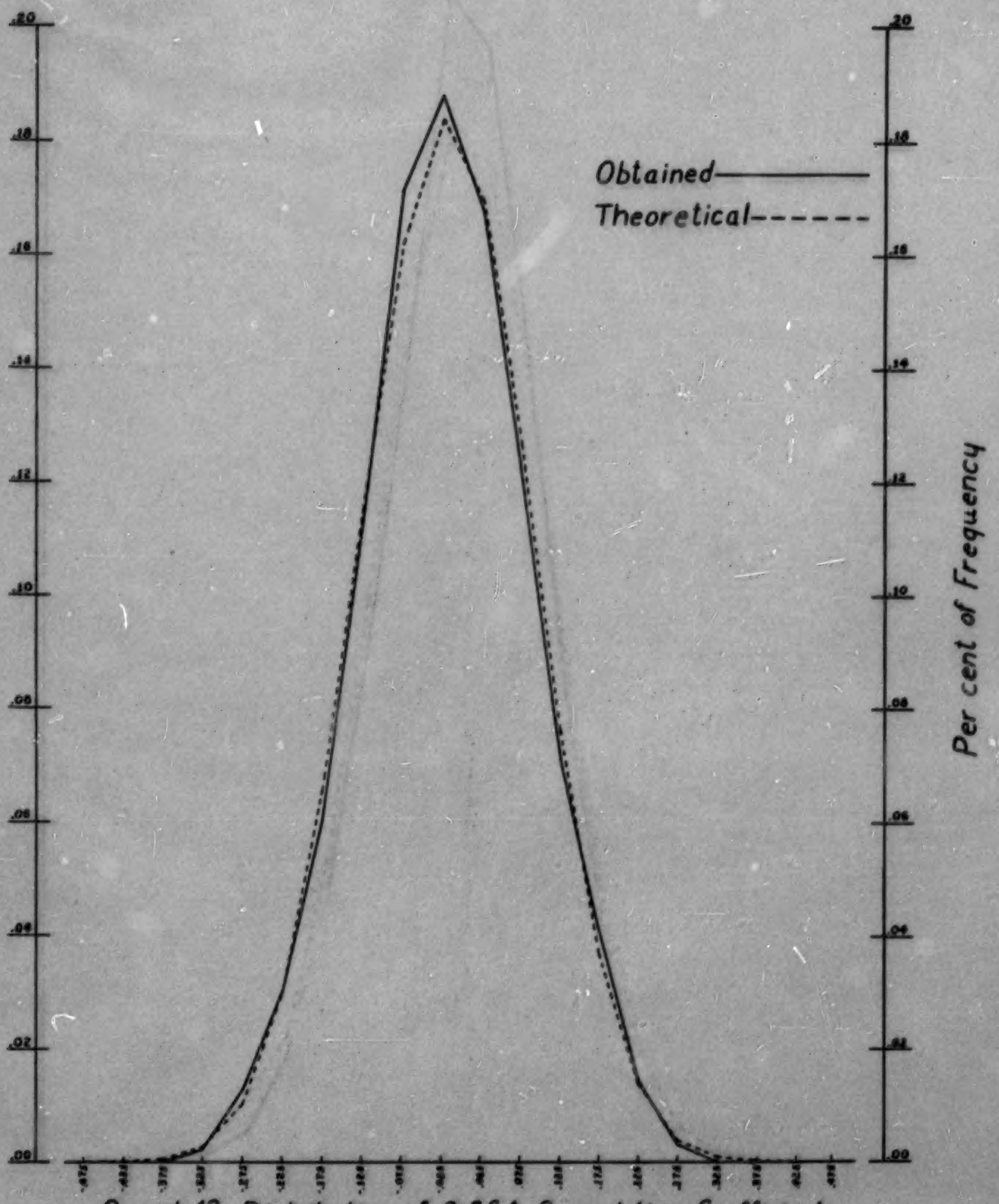
Graph 9. Distribution of 7,992 Correlation Coefficients Based on Slightly Skewed Population (p=-.010, N=30)



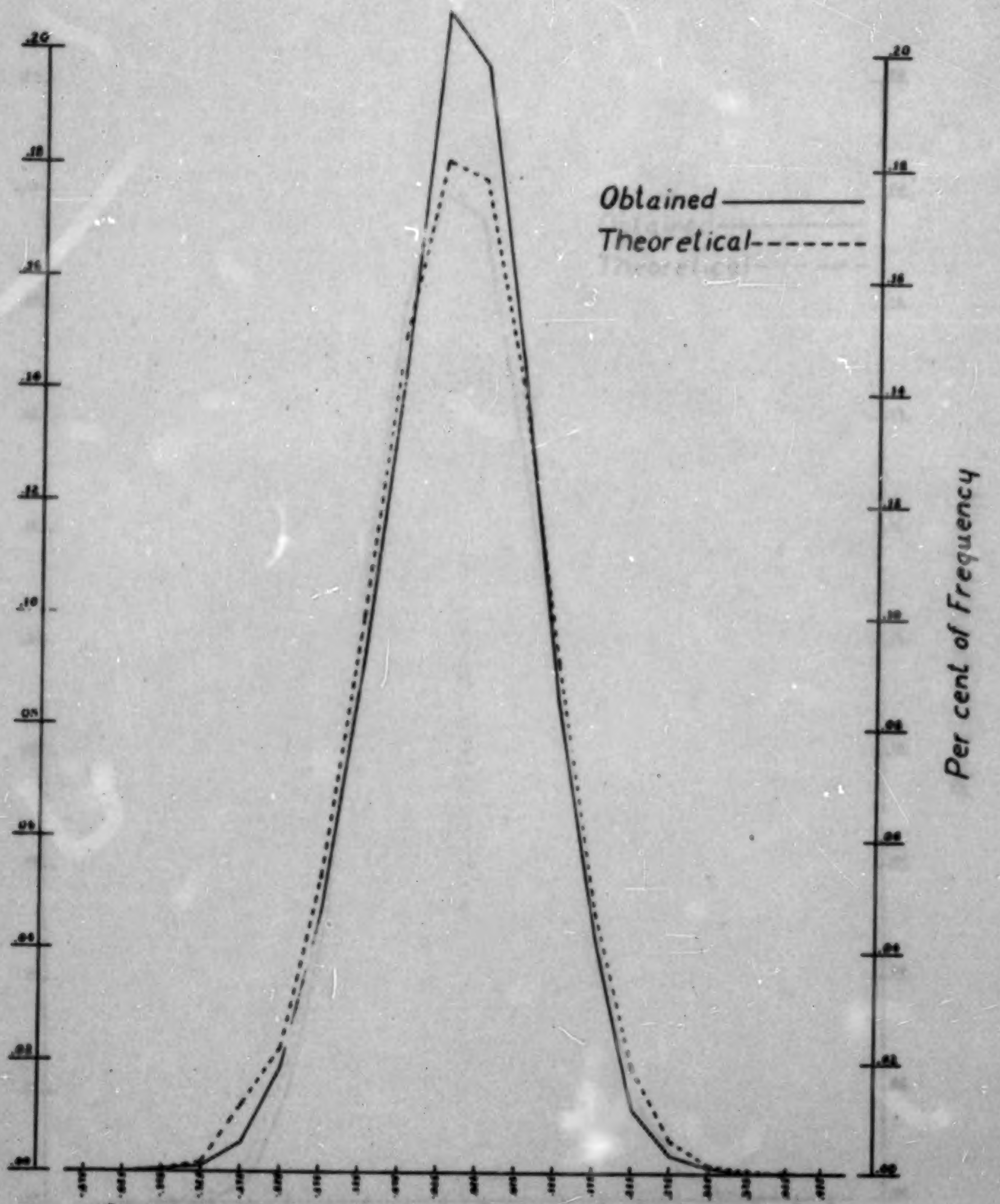
Graph 10. Distribution of 7,992 Correlation Coefficients Based on Markedly Skewed Population ($\rho = -.012$, $N = 30$)



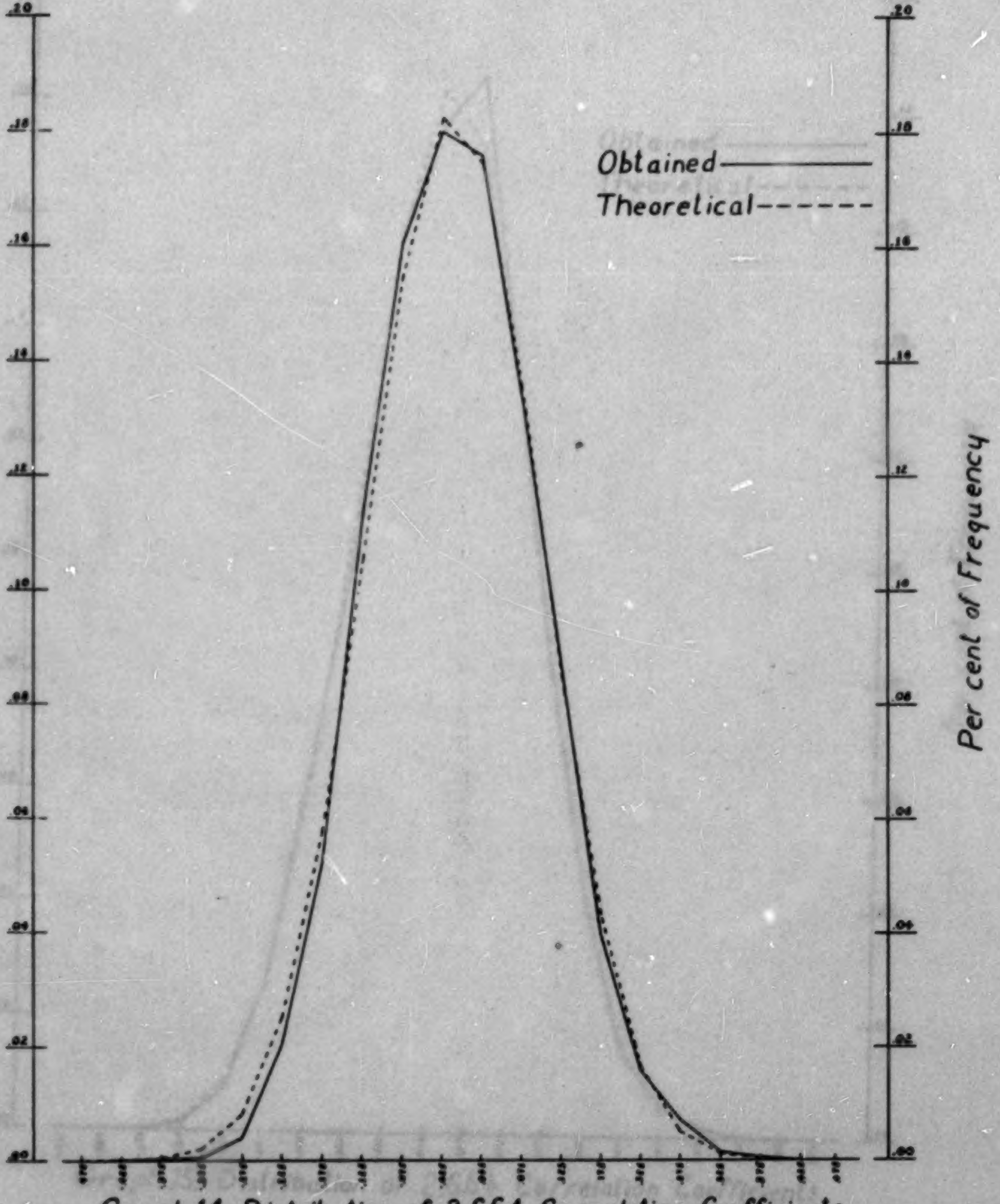
Graph 11. Distribution of 2,664 Correlation Coefficients Based on Normal Population (p = +.016, N = 90)



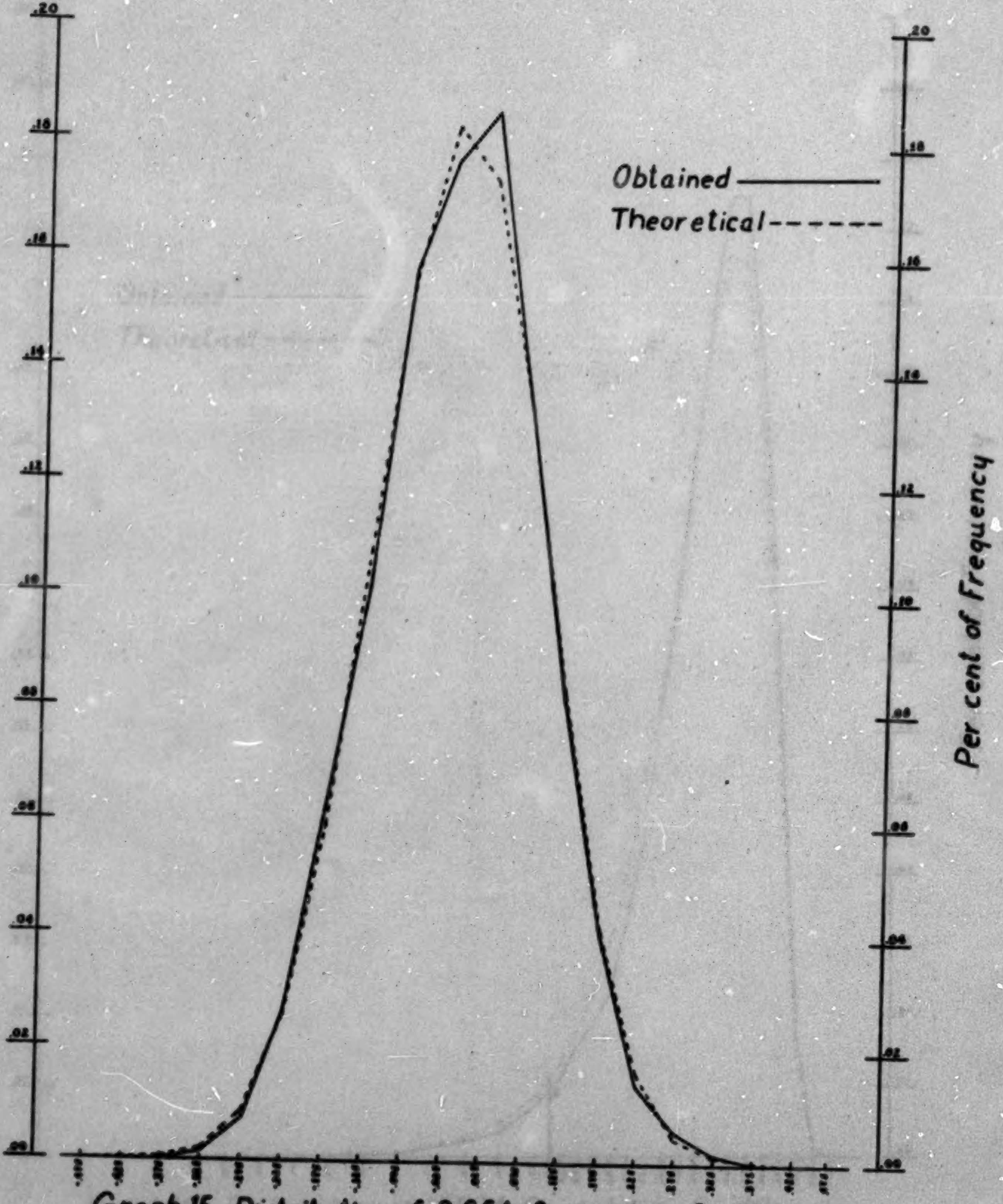
Graph 12. Distribution of 2,664 Correlation Coefficients Based on Rectangular Population ($p = -.019$, $N = 90$)



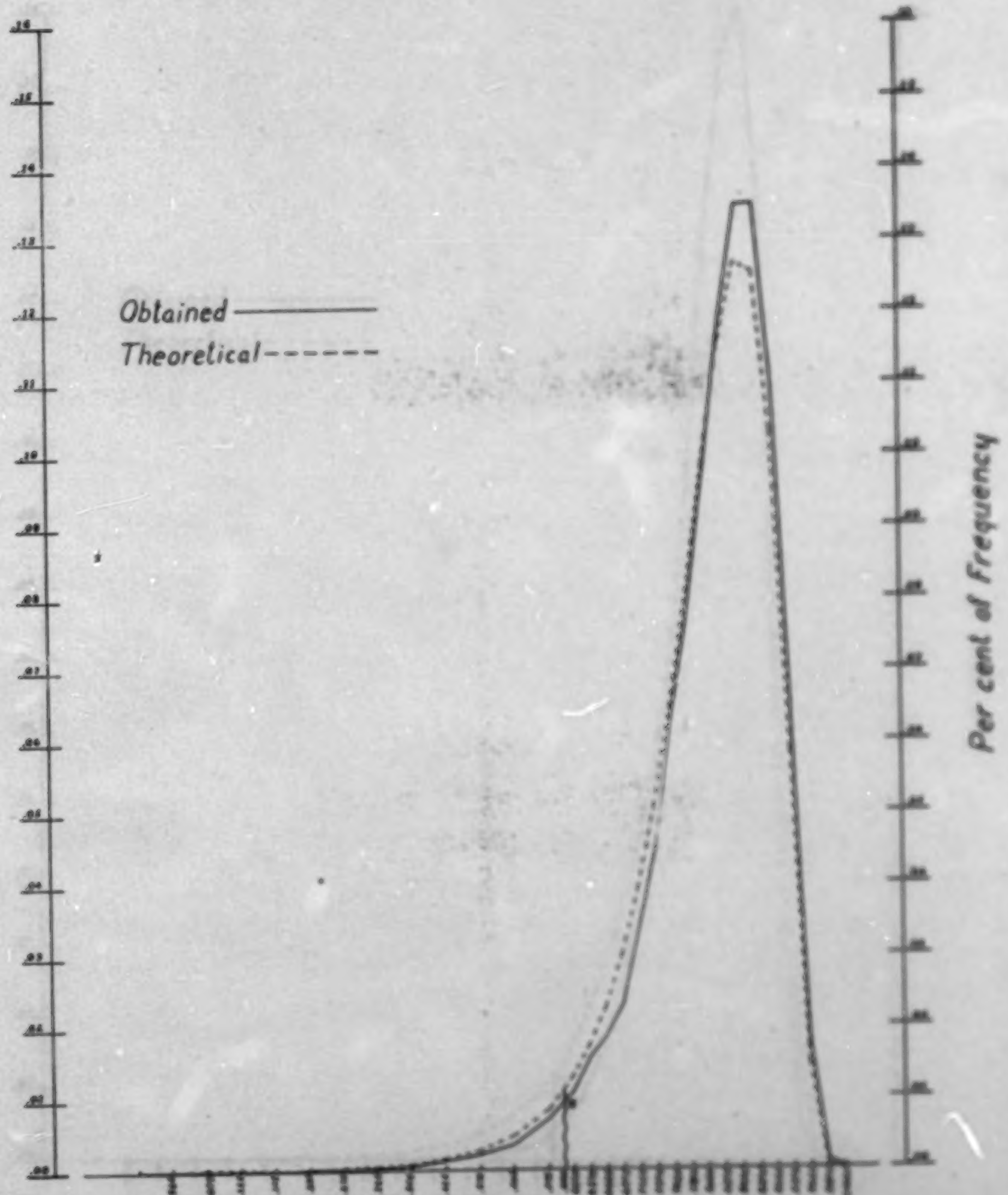
Graph 13. Distribution of 2,664 Correlation Coefficients Based on Leptokurtic Population ($p = -.010$, $N = 90$)



Graph 14. Distribution of 2,664 Correlation Coefficients Based on Slightly Skewed Population ($\rho = -.004$, $N = 90$)

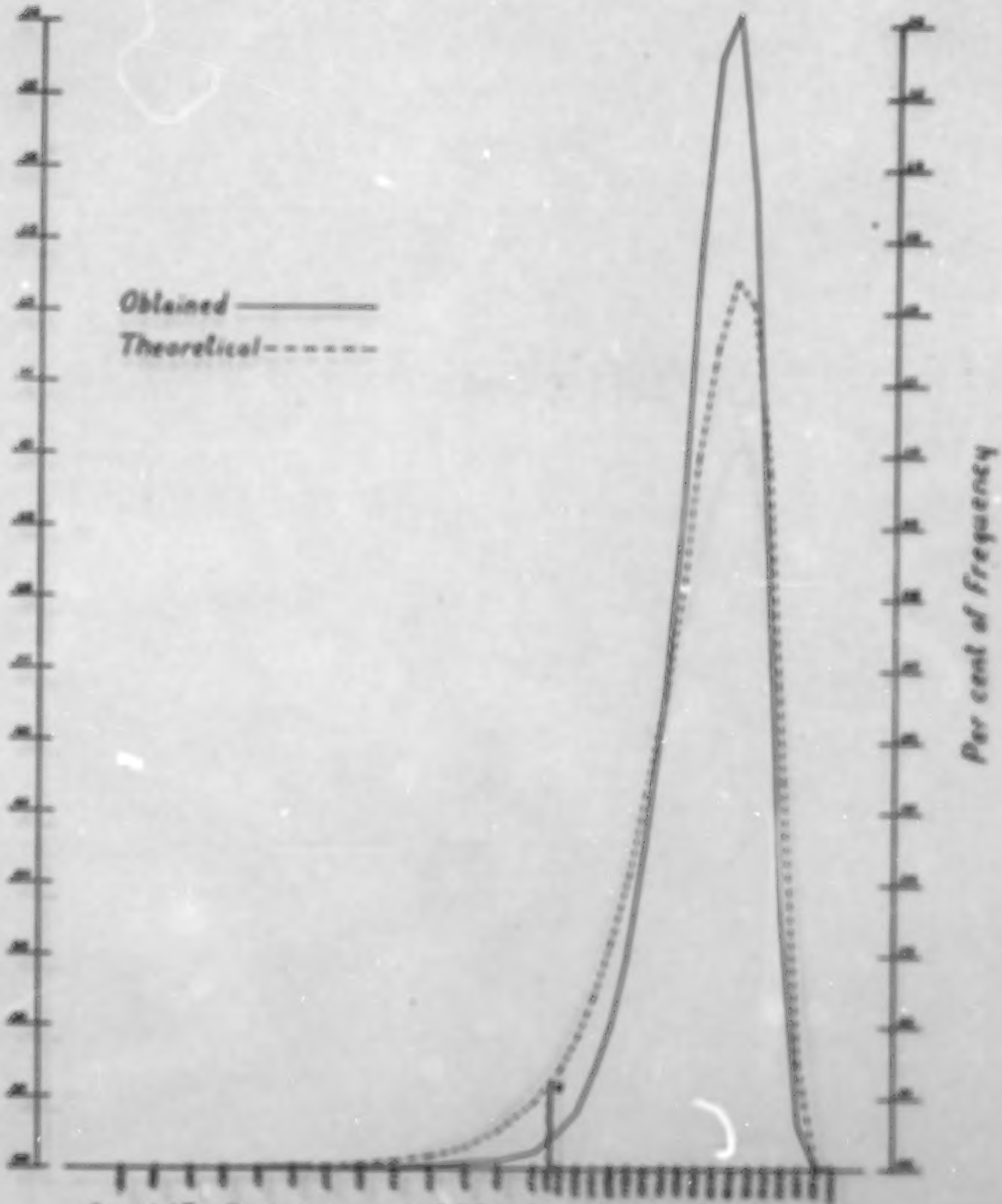


Graph 15. Distribution of 2,664 Correlation Coefficients Based on Markedly Skewed Population (p=-.012, N=90)



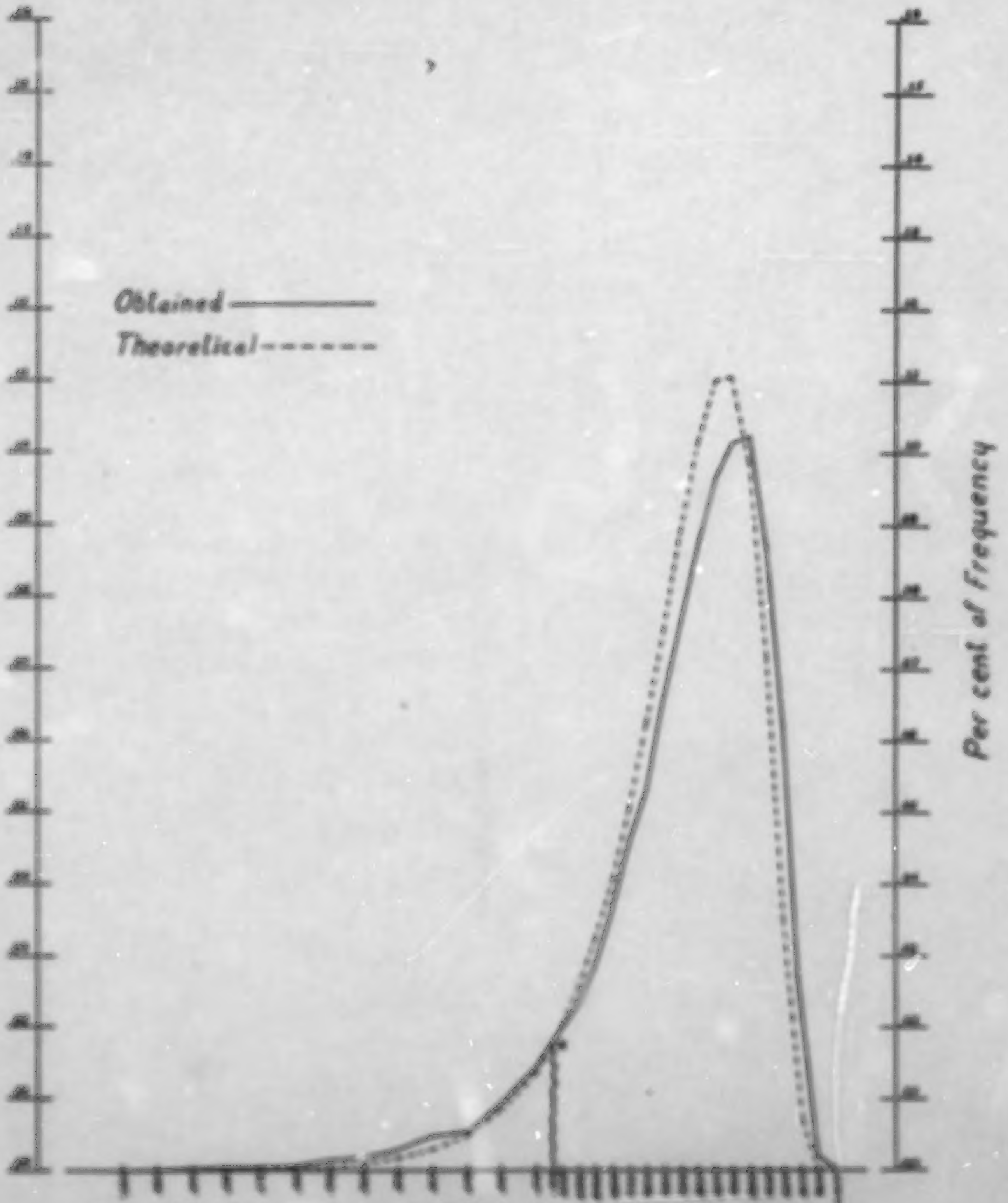
Graph 16. Distribution of 15,984 Correlation Coefficients Based on Normal Population ($\rho = .835, N = 15$)

*Per cent of Frequency halved for double width intervals.



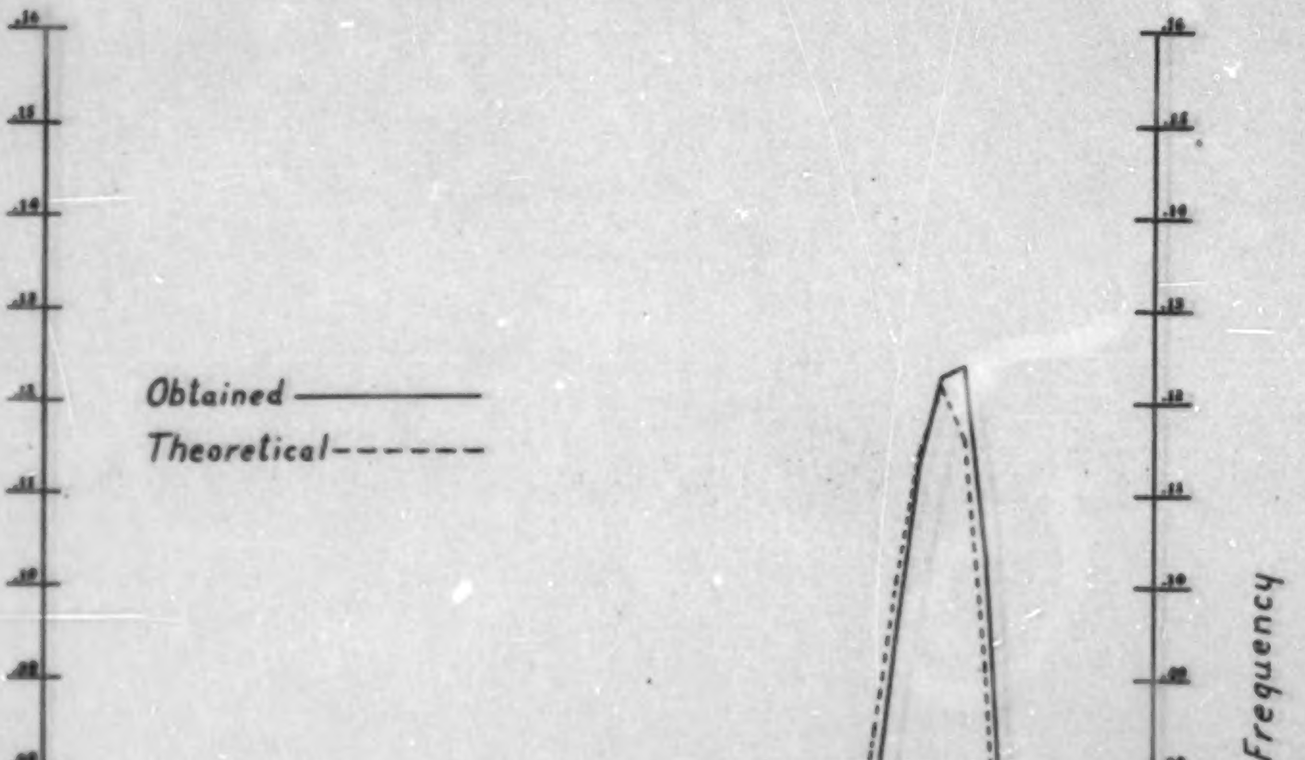
Graph 17. Distribution of 15,984 Correlation Coefficients Based on Rectangular [illegible]ulation (ρ = .831, N = 15)

*Per cent of Frequ[illegible]ted for double width intervals



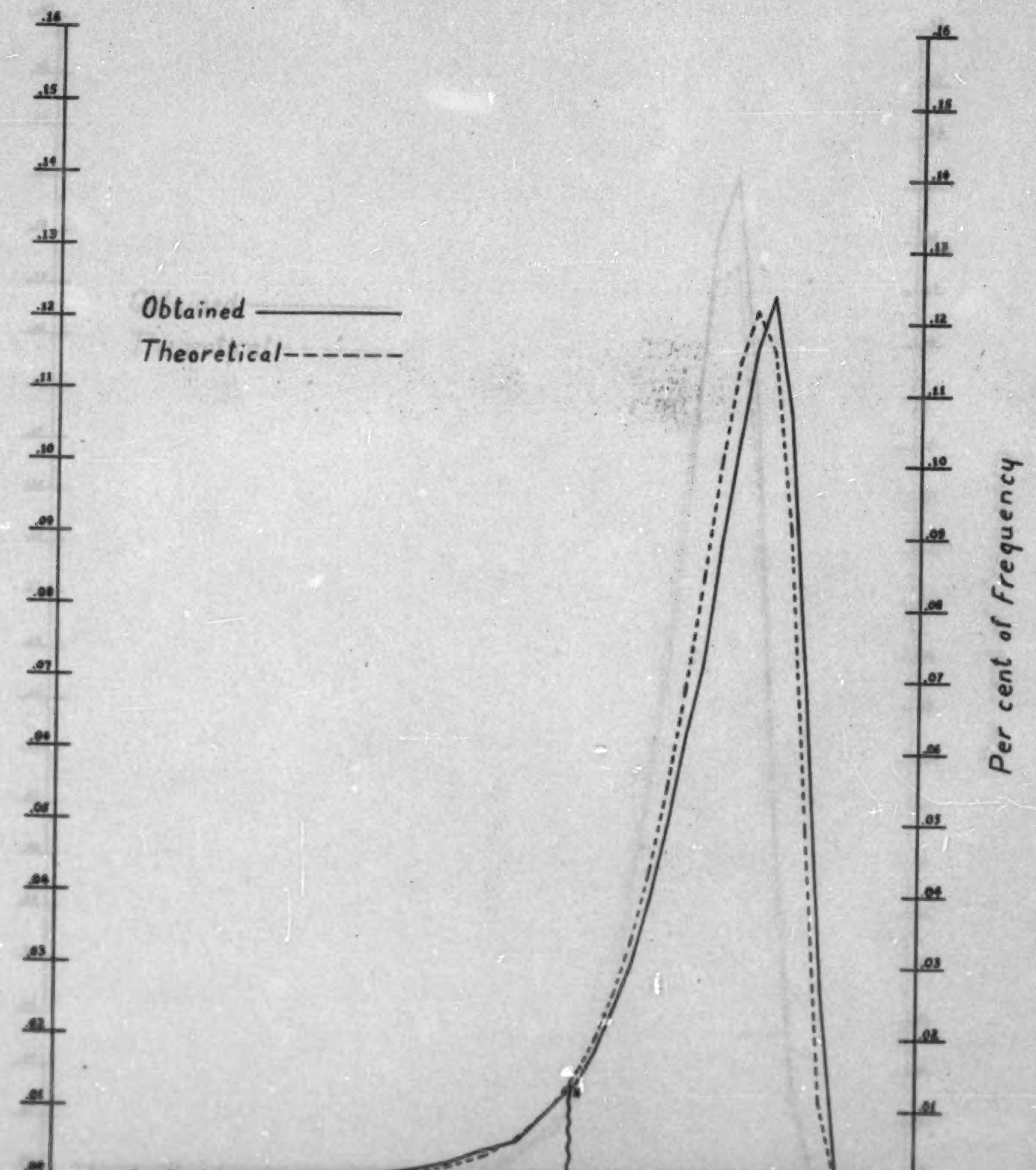
Graph 18. Distribution of 15,984 Correlation Coefficients Based on Leptokurtic Population (ρ = .809, N = 15)

*Per cent of Frequency halved for double width intervals.



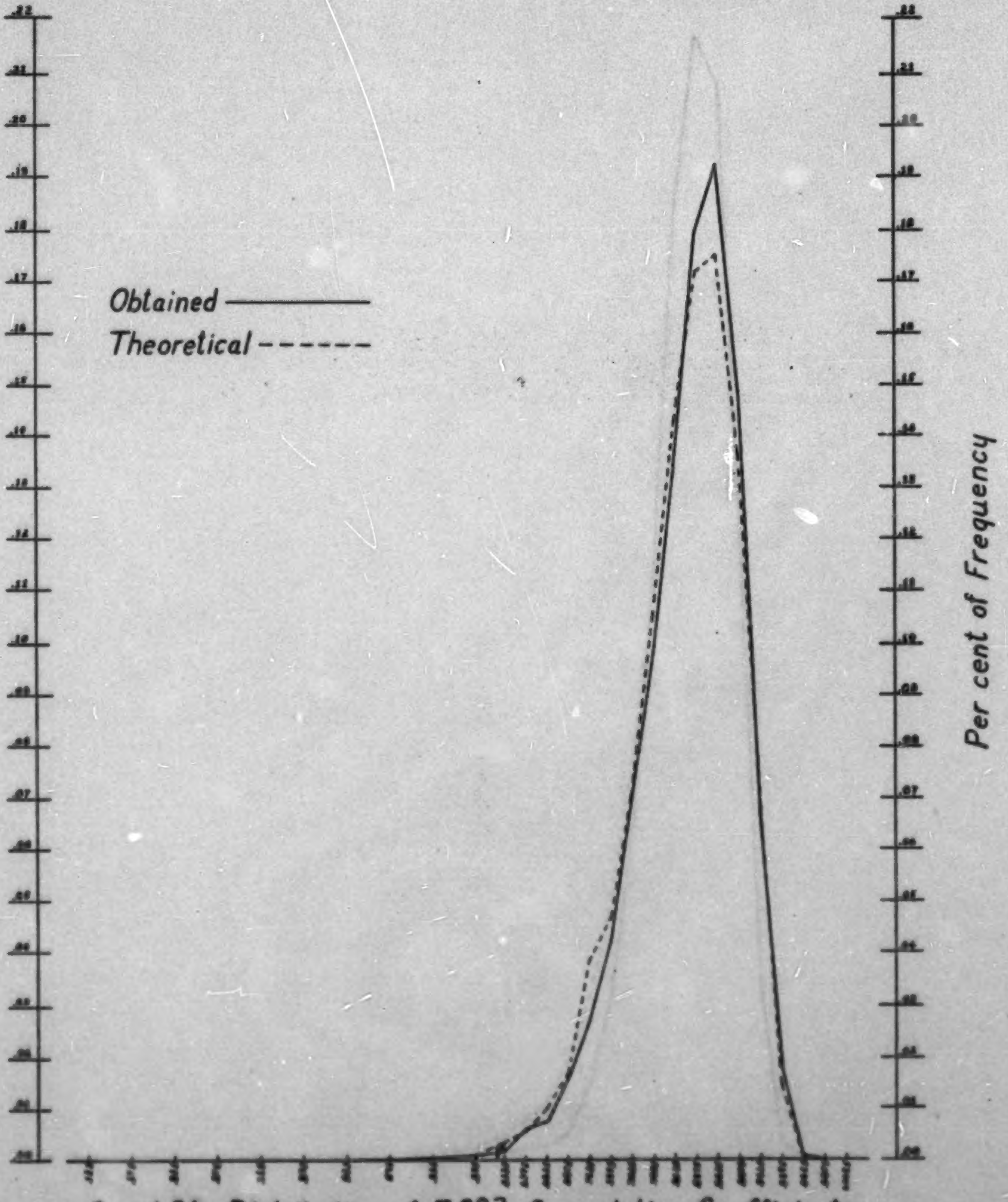
Graph 19. Distribution of 15,984 Correlation Coefficients Based on Slightlly Skewed Population (ρ=.826, N=15)

*Per cent of Frequency halved for double width intervals.



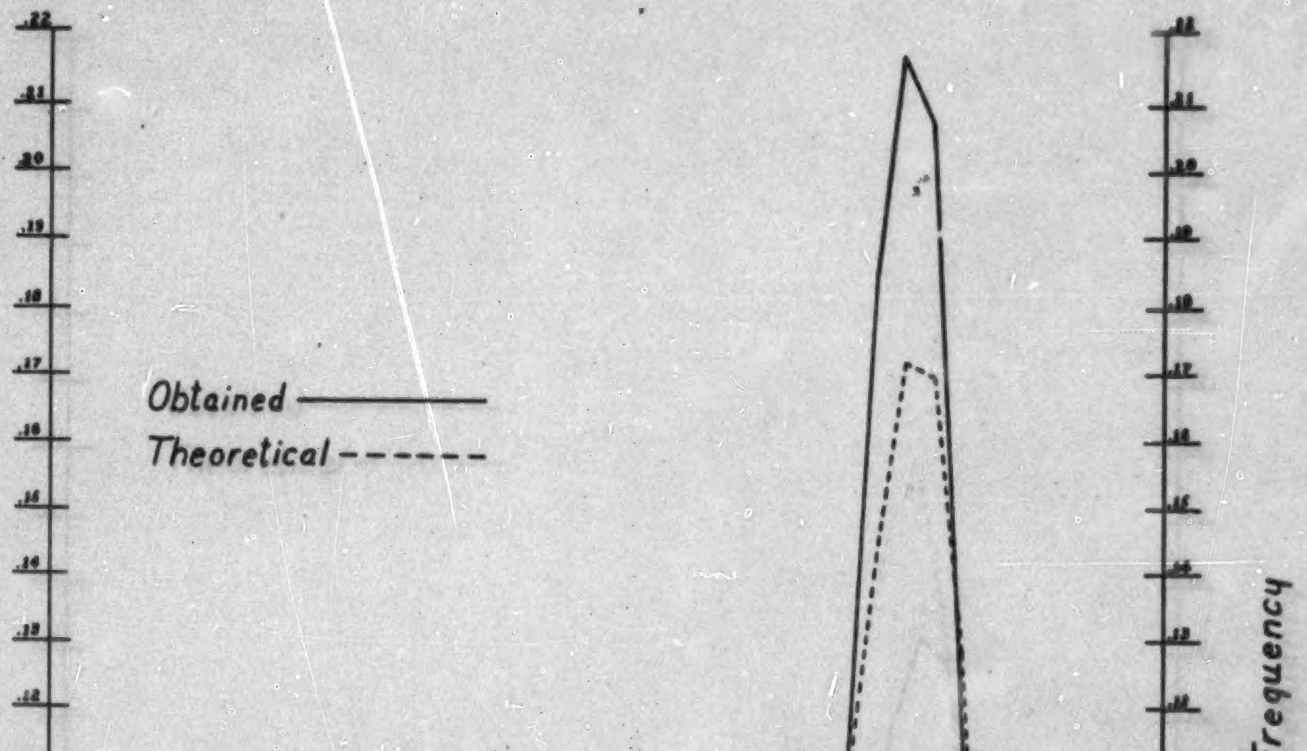
Graph 20. Distribution of 15,984 Correlation Coefficients Based on Markedly Skewed Population ($\rho = +.826$. N = 15)

*Per cent of Frequency halved for double width intervals.



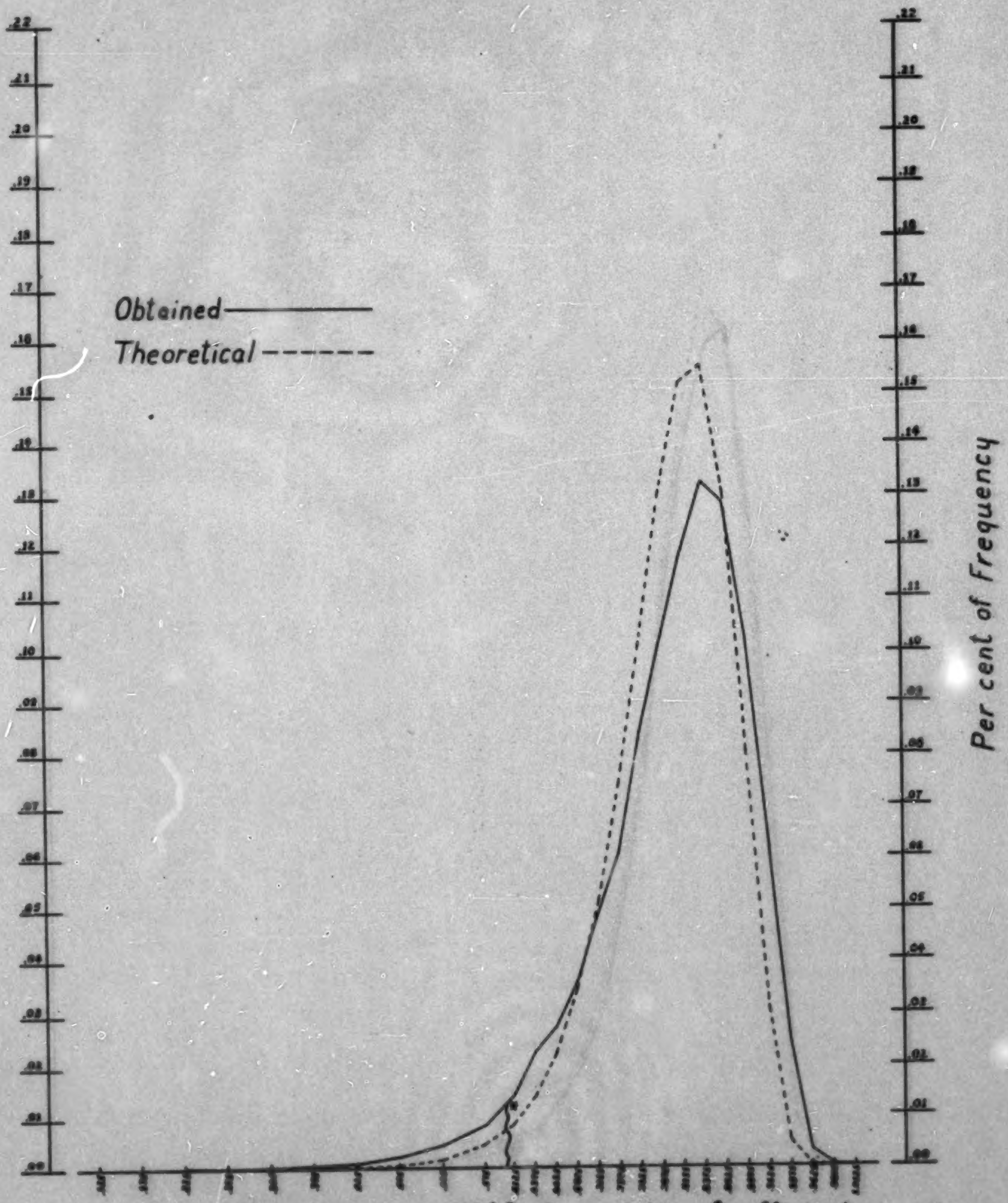
Graph 21. Distribution of 7,992 Correlation Coefficients Based on Normal Population ($\rho = +.835, N = 30$)

*Per cent of Frequency halved for double width intervals.



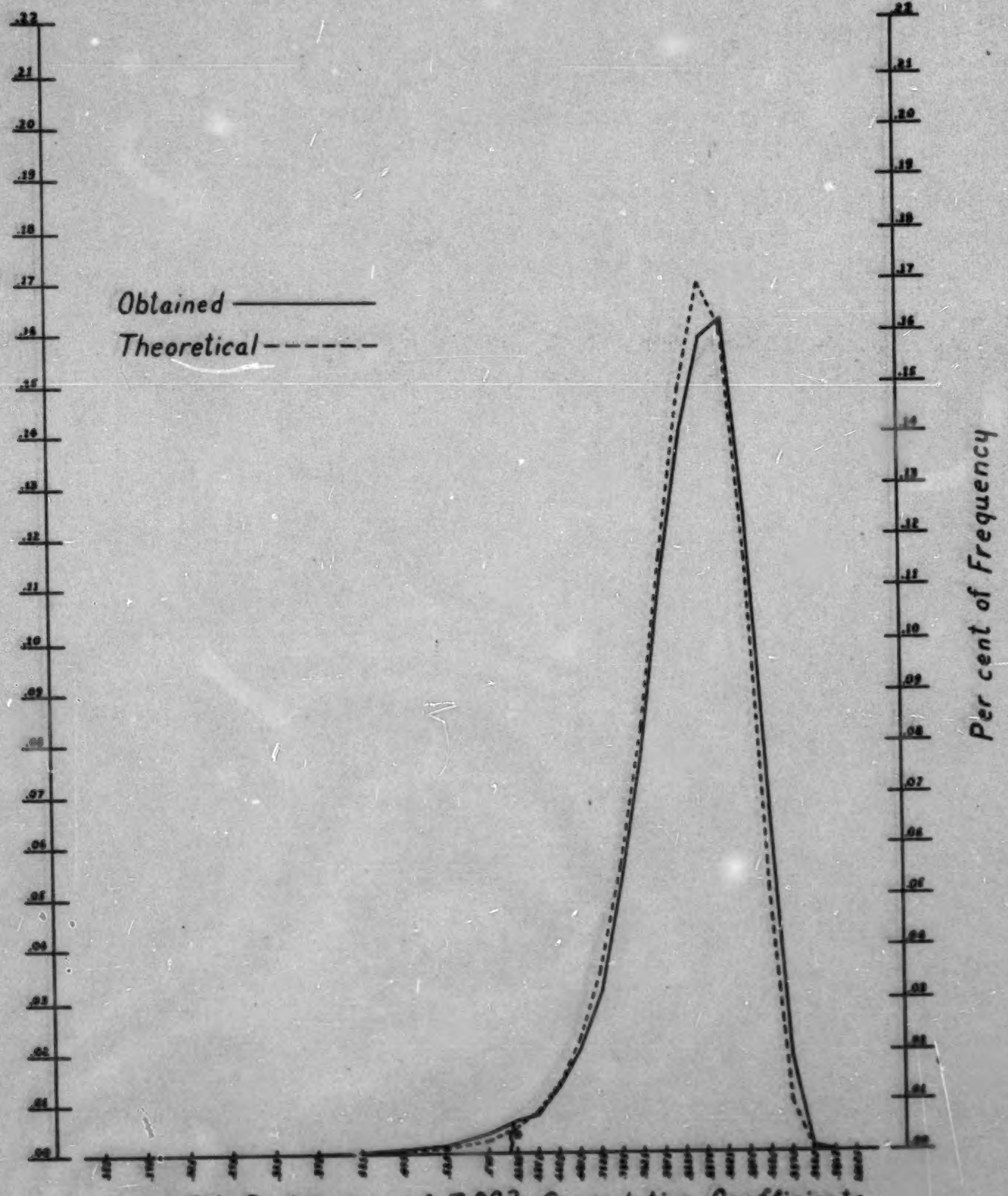
Graph 22. Distribution of 7,992 Correlation Coefficients Based on Rectangular Population (ρ = +.831, N = 30)

*Per cent of Frequency halved for double width intervals.



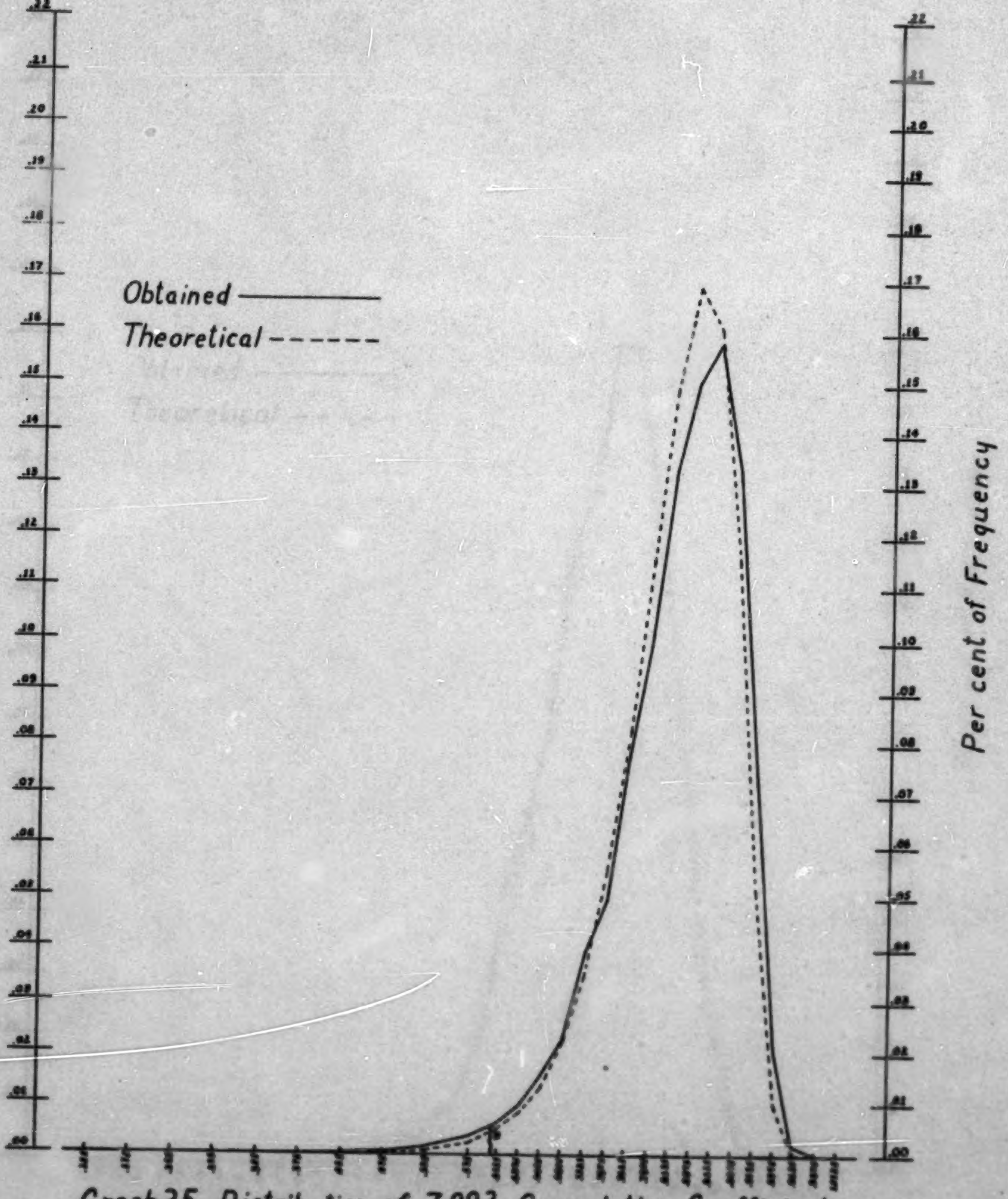
Graph 23. Distribution of 7,992 Correlation Coefficients Based on Leptokurtic Population ($\rho = +.809$. $N = 30$)

*Per cent of Frequency halved for double width intervals.



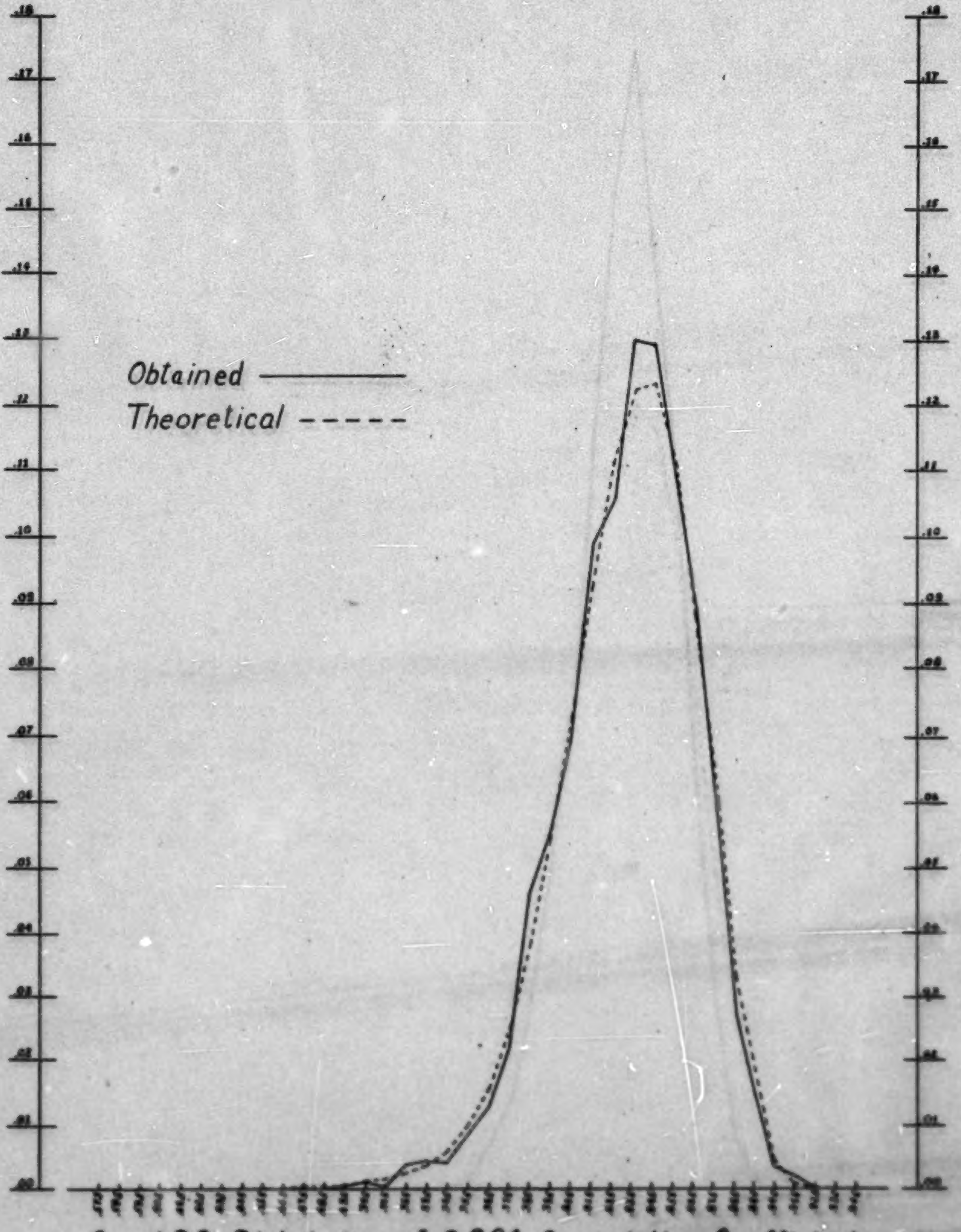
Graph 24. Distribution of 7,992 Correlation Coefficients Based on Slightly Skewed Population (ρ*=.825, N=30)

*Per cent of Frequency halved for double width intervals.

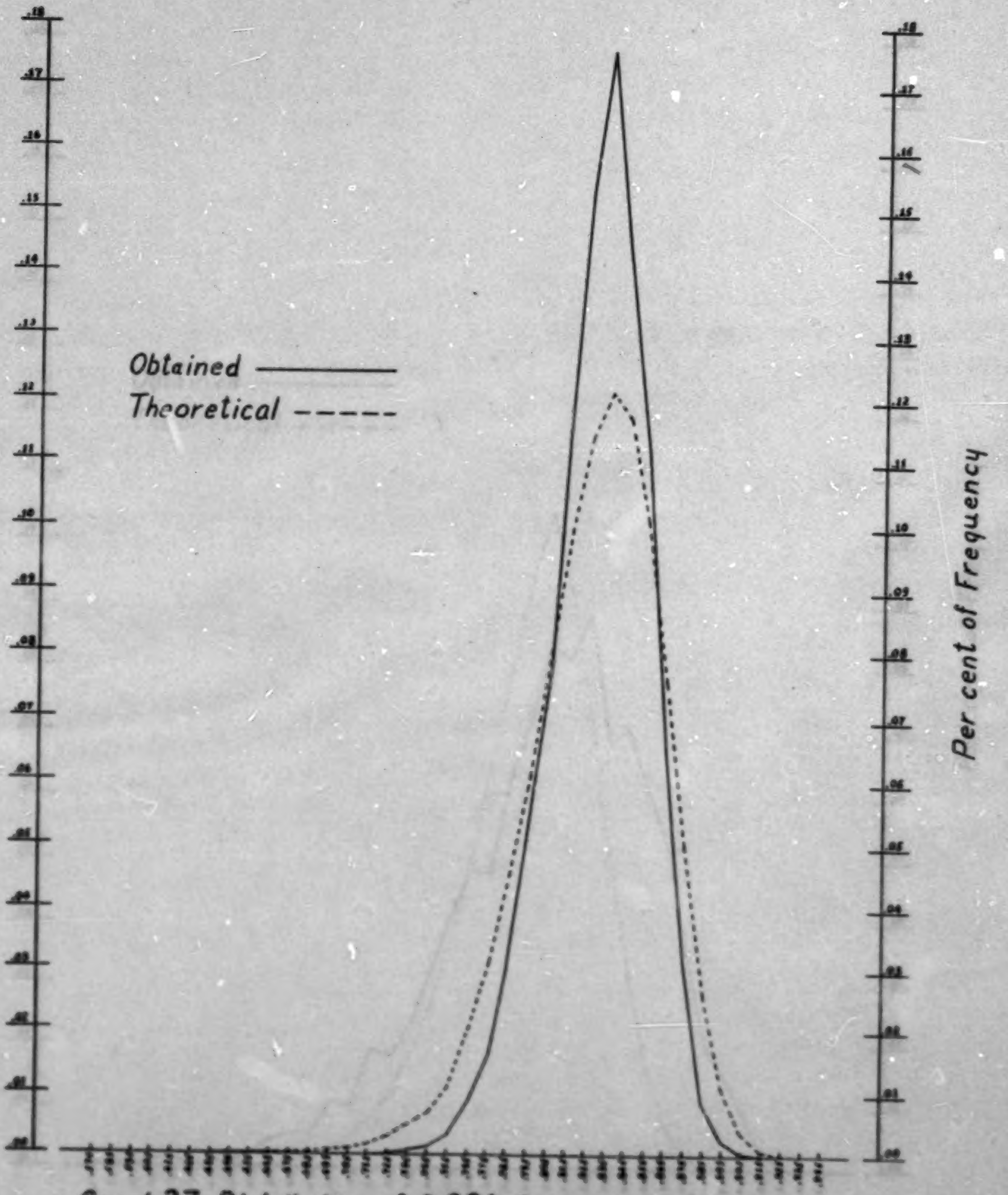


Graph 25. Distribution of 7,992 Correlation Coefficients Based on Markedly Skewed Population ($\rho = +.826$, $N = 30$)

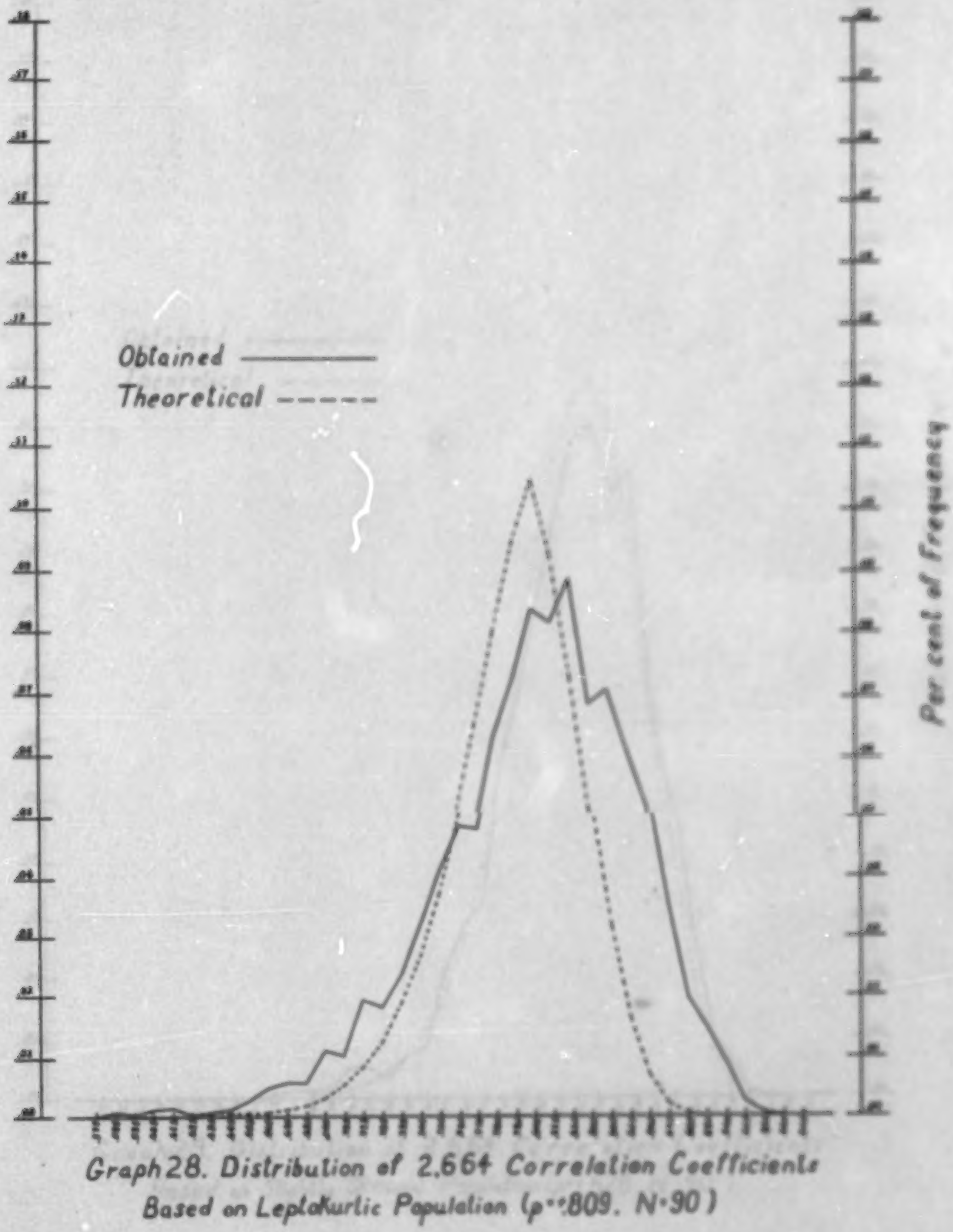
*Per cent of Frequency halved for double width intervals.



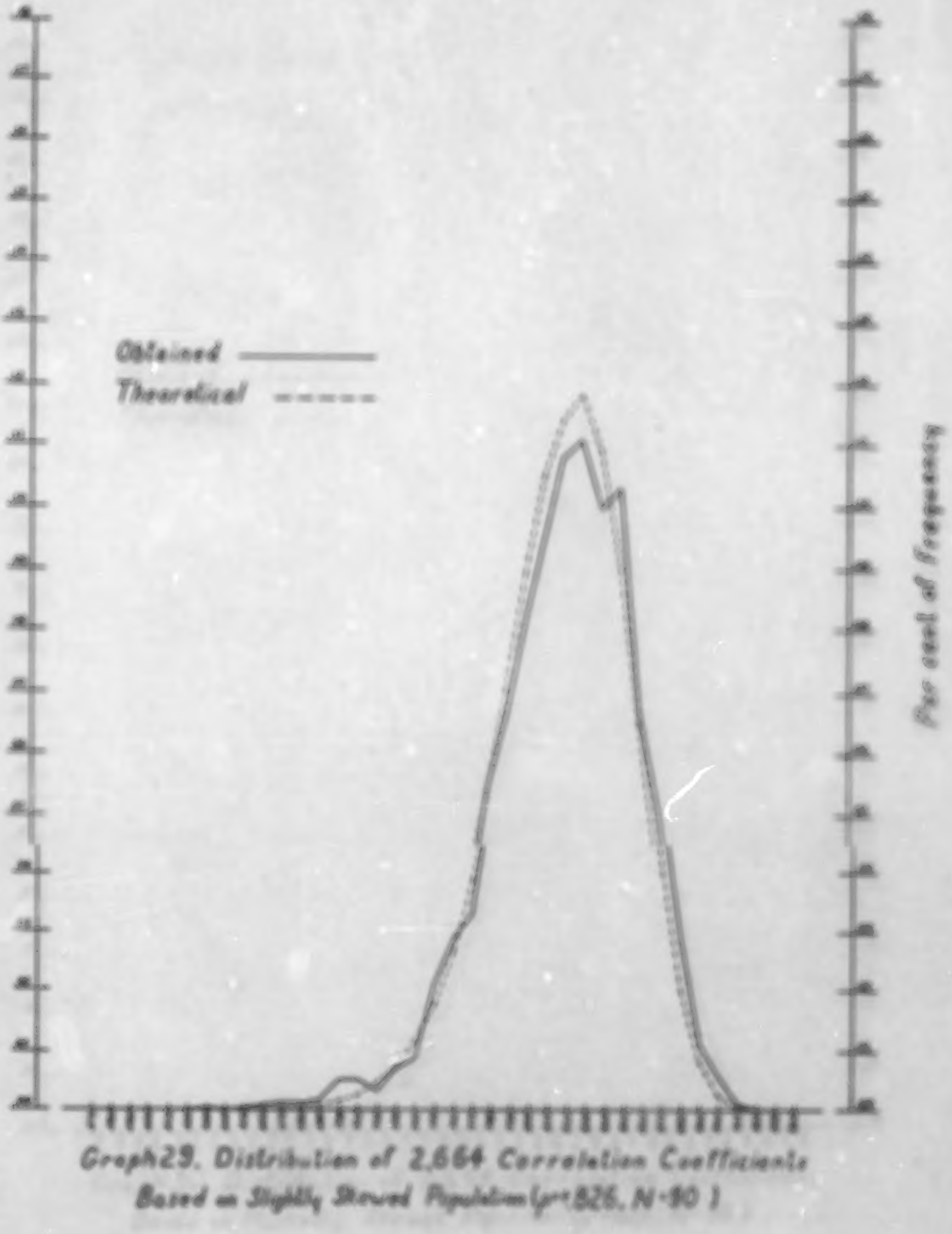
Graph 26. Distribution of 2,664 Correlation Coefficients Based on Normal Population ($\rho = .835$, $N = 90$)



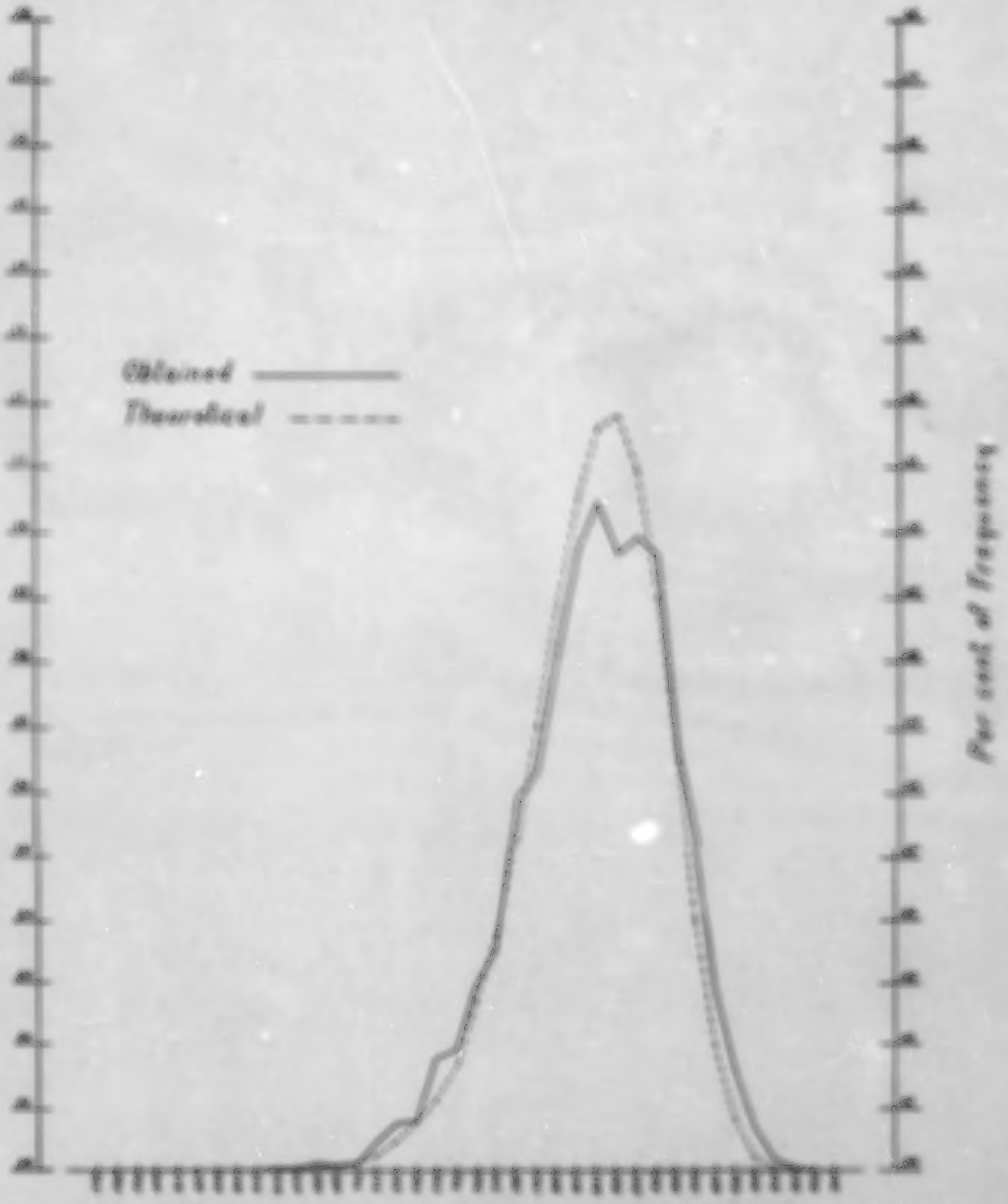
Graph 27. Distribution of 2,664 Correlation Coefficients Based on Rectangular Population ($\rho = .831$, $N = 90$)



Graph 28. Distribution of 2,664 Correlation Coefficients Based on Leptokurtic Population (ρ=.809. N=90)



Graph 29. Distribution of 2,664 Correlation Coefficients Based on Slightly Skewed Population (ρ=.826, N=90)



Graph 30. Distribution of 2,664 Correlation Coefficients
Based on Markedly Skewed Population (ρ=.826, N=30)